Scientometrics (2023) 128:2349–2369
https://doi.org/10.1007/s11192-023-04643-3

Impact of the number and rank of coauthors on *h*-index and *π*-index. The part-impact method

Péter Vinkler[1]

Received: 11 June 2022 / Accepted: 20 January 2023 / Published online: 11 February 2023

Abstract
The publication activity of 20 Price-medallists was analysed by calculating several scientometric indices from data of elite sets of their publications. The following elite sets (i.e. most cited papers within the total) were obtained: π-set (where the number of papers is equal to $\sqrt{P}$, and P is the number of total papers), h-set (in which the number of papers equals to the Hirsch index), 2π-set, $2h$-set, and the set of papers (ECP) which are cited more frequently than the average. The percentage share of papers and citations was found to increase in the mentioned rank of the elite sets, except ECP-set. The number of publications and citations in the elite sets was calculated also by different part-impact methods for *sharing credit among the coauthors.* The applied methods take into account the number or number and rank of coauthors in the by-line of the papers. It was demostrated that any of the methods changes both π and h-index significantly related to the value calculated by attributing total credit to the evaluated individual. The changes strongly depend on the number of coauthors and rank of the studied author in the by-line of the papers. Consequently, in assessing personal contribution of scientists by scientometric indices, sharing credit of publications and citations among the coauthors of publications seems to be highly relevant. Selecting the appropriate impact sharing method depends on the purpose of the evaluation.

Keywords Assessment of publications · Elite set · h-index · π-index · Sharing citations among coauthors · Part-impact method

Introduction

Distribution of credit among the coauthors of publications represents one of the main problems in assessing publications of individuals. The most simple and popular solution to the problem is the application of the *partial authorship* method (fractional allocation, see Kolthun & Hafner, 2021). Accordingly, the credit of a journal paper ($p_i = 1$) published and citations (c_i) received to the paper are equally distributed among the coauthors (a_i), i.e.: $1/a_i$ and c_i/a_i, respectively.

✉ Péter Vinkler
vinkler.peter@ttk.hu

[1] Institute of Materials and Environmental Chemistry, Research Centre for Natural Sciences, P. O. Box 286, Magyar Tudósok körútja 2, 1519 Budapest, Hungary

Although numerous sophisticated methods have been suggested in the literature (e.g. Ioannidis et al., 2019; Kolthun & Hafner, 2021; Lukovits & Vinkler, 1995; Prathap, 2021; De Solla Price & Beaver, 1966; Sangwall, 2022; Vinkler, 1993), it does not exist a widely accepted procedure for distributing the share of credit of publications and citations for the coauthors. Even the possibility to find a standard method has been questioned (Osório, 2018).

Indicators applied in scientometric evaluations of individuals, organizations or countries have been generally calculated from publication and citation data of *total* publication sets. However, after the introduction of the Hirsch-index (Hirsch, 2005), indicators referring to papers in the "core" (or elite) subsets have been preferably used. Oberesch and Groppe (2017) published a comprehensive survey on the variations of *h*-index. The paper deals also with possible methods for fractional allocation of citations in calculating the *h*-index.

The central idea behind the use of core (elite) set indicators is the assumption that indices derived from the most cited publications *may characterize the relevant publication impact of the total set more appropriately* than the mean indices of total sets (Jin et al., 2007; Leydesdorff, 2012; Schreiber, 2010; Vinkler, 2017b).

The idea for applying core journal papers in assessments finds support by the well-known fact that the distribution of citations over papers is *skewed* (Seglen, 1992). In addition, it is well known that scientific progress is advanced primarily through information in publications acknowledged by relatively high number of citations (Aksnes, 2003; Plomp, 1990; Vinkler, 2010a, 2017a). Consequently, publications with high influence may be revealed by determining relatively highly cited journal papers.

Core (elite) subsets can be obtained by different statistics: *h*-statistics, *g*-statistics, *percentage* statistics, *π*-statistics, *πv*-statistics, etc. The core or elite part of publication sets is termed as *h-core*, *g-core*, *π-core*, or *top-1%*, *10%*, etc. of total, according to the method applied. Naturally, the publications in the sets studied are ranked according to the decreasing citation frequency by each method.

The preferred use of elite set indicators has resulted in a plethora of scientometric impact indices (Leydesdorff, 2012; Schreiber, 2010; Schreiber et al., 2012; Todeschini & Baccini, 2016).

The impact indices derived through the mentioned statistics, e.g.: *h*-index (Hirsch, 2005) and *h*-related indices (Bornmann et al., 2011), *g*-index (Egghe, 2006), *π*-index (Vinkler, 2009), *πv*-index (Vinkler, 2010b), *I3*-index, indices of percentiles and percentile classes, *PR*(*2*, *10*) or *PR*(*2*, *50*) or *PR*(*6*) (Bornmann & Marx, 2014; Bornmann et al., 2013; Leydesdorff, 2012), top-1% highly-cited papers (Wagner, Lin Zhang., & Leydesdorff, 2022), etc. strongly depend on the *bibliometric features* of the corresponding field.

There are some methods applying *outside standards* for obtaining elite subsets. One of these methods takes those publications as "elite" to which the number of citations obtained is *higher than the mean citation rate of the corresponding field* (Vinkler, 2017a). According to the most wide spread method, the journal papers (e.g. of a country) in the top-1% of highly cited papers in the corresponding field may be accepted as most influential on world level (see e.g. WoS Essential Science Indicators).

Kolthun and Hafner (2021) tried to verify the *h*-index by comparing the appropriate values of individuals with their scientific awards. They conclude: „…*fractional allocation* (of citations) *improves the effectiveness and predictive power of research metrics, and h-frac is consistently the most reliable bibliometric indicator. Our results suggest that the* (recent) *use of the h-index in ranking scientists should be reconsidered…*".

The publications in elite subsets of a scientific field may represent *hot topics* or relevant *core information of the field* depending on the time period applied. Science is developing permanently. Therefore, the dynamic study of the publications in elite subsets is highly relevant (Egghe, 2007).

The present paper compares the *h-index* and *π-index* calculated by attributing total credit (both publication and citation) to one of the coauthors (i.e. to the evaluated scientist) of multiauthored papers with *h-index* and *π-index* obtained by attributing only the *relevant part* of papers and citations to the scientist studied. The methods calculating impact indices from data of papers (and citations) in the *relevant part* of multiauthored publications may be called as *part-impact* methods. Accordingly, the research tasks of the present publication may be summarized as follows:

- Study of impact of the number and rank of authors in the byline of publications on *π*-index and *h*-index applying Full Authorship Method (FAM), Partial Authorship Method (PAM), Practical Rank Score Method (PRSM), and Single plus First Authorship Method (SFAM). (For FAM, PAM, PRSM, and SFAM, see Data and methods.)
- Comparison of the percentage share of journal papers and citations in the following elite subsets: *π*-set, *h*-set, 2*π*-set, 2*h*-set.
- Revealing the difference in the mean citation rate (C/P) of papers in the elite subsets (*π*, 2*π*, *h*, and 2*h*-set).
- Study of the application of part-impact methods on the rank of scientists according to *h*-index and *π*-index.

Data and methods

For investigating the impact of the number of coauthors of papers of a scientist on *h*-index and *π*-index, total sets of journal papers published by 20 Price medallists were collected. The persons selected (T Braun, B Cronin, L Egghe, W Glänzel, P Ingwersen, L Leydesdorff, B R Martin, K W McCain, H F Moed, F Narin, O Persson, R Rousseau, A Schubert, H Small, M Thelwall, AFJ vanRaan, P Vinkler, L Waltmann, H White, M Zitt) may be regarded as scientists with outstanding contribution to the field of quantitative studies of science.

The present study does not intend to evaluate the publication activity of the mentioned scientists. The publication and citation data of the studied researchers are taken as model for investigating features of some elite set indicators calculated with part-impact methods.

The publication and citation data of the mentioned scientists were collected from Web of Science in Mai, 2021. Only papers classified by WoS as: *Information Science Library Science* and *Business Economics* were selected (ILBE papers). Each ILBE paper was reviewed and only those belonging to scientometrics, bibliometrics or informetrics were taken into account (TP papers). The total number of TP publications: 2718, to which 99,396 citations were arrived up to the mentioned time.

The following indicators were calculated for each scientist: total number of papers (TP), total number of citations (TC) to TP papers, mean citation rate, TC/TP, number and per cent of papers cited more frequently than the mean (ECP and ECP%, resp.), per cent of papers in the elite subsets (*π*, 2*π*, *h*, 2*h*) and that in the rest of total [(TP-P(2h)]%. The number of papers in the *π*-set = $\sqrt{TP}$ (Vinkler, 2009). The number of papers in the *h*-set is equal to the *h*-index. The number of papers in the 2*h* or 2*π* set is equal to two times the number of papers in the *h* or *π* set, respectively.

The number and rank of coauthors of TP papers was taken from WoS. Sharing citations among the coauthors was made by the full credit method and by three part-impact methods:

- *Full Authorship Method* (FAM): total number of citations received by a journal paper is given to the author studied independent of number of coauthors and rank of the assessed author in the byline of the paper.
- *Partial Authorship Method* (PAM): the studied author is given c_i/a_i share of citations to the paper (c_i), where a_i is the number of authors of the i-th paper.
- *Practical Rank Score Method* (PRSM). According to this method, the number of authors and rank of the studied author in the byline of the paper is taken into account in calculating the share in citations of the studied author (Table 1).

The citations of the author studied will be obtained as a product of the total number of the citations to the paper (c_i) multiplied by the corresponding *cooperation factor* (Table 1). The cooperation factors applied are calculated according to Vinkler (1993) and Lukovics and Vinkler (1995). (E.g. $c_i = 30$ to a paper with $a_i = 3$ authors where the scientist studied is the 3rd author. Accordingly, the corresponding share: $30 \times 0.20 = 6$, see Tables 1, 2).

The authors ranked higher than sixth, each will be given a factor of 0.05. This way the sum of the cooperation factors of the authors would be higher than unity. Therefore, the factors of the individual authors should be normalized into unity. In case of e.g. 8 authors: the sum of the cooperation factors $= 1 + 2 \times 0.05 = 1.1$. After normalization, the first, second, etc. author will obtain $0.35/1.1 = 0.318$; $0.25/1.1 = 0.227$; $0.15/1.1 = 0.136$; $0.1/1.1 = 0.09$; $0.1/1.1 = 0.09$; $0.05/1.1 = 0.045$; $0.05/1.1 = 0.045$; $0.05/1.1 = 0.045$ scores. The sum of scores so obtained $= 0.996$ (rounded to unity).

- *Single plus First Authorship Method* (SFAM): the papers authored only by the studied scientist, plus that where this author ranks as first, independent of the number of coauthors, are classified as SFA-papers.

It is obvious that in the PA, PRS, and SFA-set the *rank of the papers* by the number of citations will be *different* from that in the original set (Full Authorship Method, FA-set). Accordingly, also h-index, π-index, etc. obtained for the FA-set will be different from the indices calculated for the PA, PRS and SFA-set (see Table 2).

Table 1 Cooperation factors for applying Practical Rank Score Method (PRSM)

Number of authors of the publication	Rank of authors					
	1	2	3	4	5	6
	Cooperation factor					
1	1.00					
2	0.65	0.35				
3	0.55	0.25	0.20			
4	0.50	0.25	0.15	0.10		
5	0.40	0.25	0.15	0.10	0.10	
6	0.35	0.25	0.15	0.10	0.10	0.05

Table 2 Model for demonstrating the change in rank and citations to the papers on *h*-index by applying Full Authorship Method (FAM), Partial Authorship Method (PAM), and Practical Rank Score Method (PRSM)

Number of authors (a_i) of the paper	Citations (c_i) obtained to the paper	Rank of the paper by **FAM**	Modified number of citations: $mc_i = c_i/a_i$	Rank of the papers by **PAM**	Rank of the *studied author* in the by-line of the paper	Modified c_i by the number of authors and rank of the studied author (rounded)	Rank of the papers by **PRSM**
3	18	1	6	4	1	18 × 0.55 = 9.9 (10)	2
2	16	2	8	3	2	16 × 0.35 = 5.6 (6)	4
1	15	3	15	1	1	15 × 1.00 = 15	1
3	12	4	4	5	3	12 × 0.20 = 2.4 (2)	5–6
1	9	5	9	2	1	9 × 1.00 = 9	3
4	8	6	2	7	2	8 × 0.25 = 2	5–6
2	6	7	3	6	1	2 × 0.65 = 1.3 (1)	7

h-index (FAM) = **6**; *h*-index (PAM) = **4**; *h*-index (PRSM) = **4**

π-index (FAM) = 0.49; π-index (PAM) = 0.32; π-index (PRSM) = 0.34

$P(\pi) = \sqrt{7} = 2.6$ (rounded: 3)

The statistical calculations and figs. were made with TIBCO Statistica 13.4/13.5 program.

Results and discussion

Share of papers in the elite sets

The total number of journal papers (TP) published by the scientists studied, mean number (TC/TP) of citations (TC) received, percentage share of papers cited more frequently than the mean (ECP%), and that in π, 2π, h, $2h$, and (TP-2h)-set are given in Table 3. The studied scientists published 135.9 papers (SD = 105.5) in average during the investigated period (1975–2021). The dynamic range of the data is rather wide (from 27 up to 369).

The data in Table 3 show that the mean percentage rate of the papers increases in the rank: P(π), P(2π), ECP, P(h), P(2h). The mean share of papers in the h-set (32.83, SD = 14.07) is 3.13 times higher than that in the π-set (10.50, SD = 3.73). The difference between the mean P(π)% and P(h)% is *highly significant* ($p < 0.001$). The number of papers in h-set may be similar to that in π-set only in extreme cases (Vinkler, 2017a).

The ratio of papers within a given set with *higher* number of citations than the mean (Table 3) may be characteristic for the distribution of citations over papers. The mean ECP% of the studied scientists is 26.56, whereas the standard deviation of the mean is rather low (SD = 5.78). Accordingly, *the share of papers cited higher than the average seems to be similar in the publication sets of eminent scientometricians.*

The data in Table 3 also show, the mean percentage rate of papers in the π-set, P(π)% (10.50, SD = 3.73) is significantly ($p < 0.001$) *lower* than the share of papers with higher citation rate than the mean (ECP% = 26.56, SD = 5.78). At the same time, mean P(h)% *does not differ significantly* ($p = 0.073$) from the mean ECP%, whereas its value is somewhat *higher* (32.83%, SD = 14.07). The difference may be attributed to the higher number of highly cited papers in the h-core compared to π-core.

The Spearman rank-order correlation coefficients reveal significantly high and positive correlation ($r = 0.88$) between the percentage share of papers in the π and h-set, P(π)%, P(h)%, resp. It is obvious however; the greater the size of the publication set (TP), the lower the share of papers in the π and 2π-set, relatively. Therefore, the correlation coefficients are negative: $r = -1.00$ and $r = -1.00$, respectively. The coefficient between TP and P(h)% is also negative but it differs from unity: $r = -0.88$. In contrast, the share of papers in the (TP-2h)-set seems to be greater with greater number of papers ($r = 0.69$). I.e. publishing relatively many papers, it would mean that beyond a limit, the papers would receive fewer citations. These papers are published in journals with lower impact factor, in average.

It is unexpected however that *no significant correlation* (at $p < 0.05$) is found between ECP% and share of papers in the elite sets, P*(π)%*, P*(h)%*: $r = 0.14$ and $r = 0.31$, respectively. The reason for that may be the skewed distribution of citations over papers.

Share of citations in the elite sets

Indicators applying citations represent a preferred role in evaluative scientometrics. The average number of total citations obtained by the scientists in the period studied is 5038.6 (SD = 3806.1) (Table 4). The dynamic range is rather wide: 16,051 (Leydesdorff) – 1190 (Vinkler). Table 4 shows also the share of citations in the different elite subsets.

Table 3 Total number of papers (TP) published by the studied scientists, mean citation rate, TC/TP, number and per cent of papers cited more frequently than the mean (ECP and ECP%, resp.), per cent of papers in elite subsets (π, 2π, h, $2h$) of different size and that in the rest of total [(TP-P(2h)]%

Name	TP	Mean TC/TP	ECP	ECP%	P(π)%	P(2π)%	P(h)%	P(2h)%	[(TP-P(2h)]%
Braun	80	35.99	21	26.25	11.25	22.50	35.00	70.00	30.00
Cronin	278	15.16	72	25.90	6.12	12.23	11.51	23.02	76.98
Egghe	184	21.23	39	21.20	7.61	15.22	15.76	31.52	68.48
Glänzel	251	36.90	76	30.28	6.37	12.75	21.91	43.82	56.18
Ingwersen	82	32.06	16	19.51	10.98	21.95	31.71	63.42	6.59
Leydesdorff	315	50.96	87	27.62	5.71	11.42	20.32	40.64	59.37
Martin	27	51.48	10	37.04	18.52	37.04	59.26	100.00[a]	(40.74)[b]
McCain	65	43.51	12	18.46	12.31	24.62	33.85	67.70	32.31
Moed	111	52.93	31	27.93	9.91	19.82	37.84	75.68	24.32
Narin	50	103.08	12	24.00	14.00	28.00	56.00	100.00[a]	(44.00)[b]
Person	46	46.52	10	21.74	15.22	30.43	45.65	91.30	8.70
Rousseau	281	18.61	64	22.78	6.05	12.10	12.46	24.92	75.09
Schubert	150	30.83	44	29.33	8.00	16.00	24.00	48.00	52.00
Small	64	46.64	20	31.25	12.50	25.00	40.63	81.26	18.75
Thelwall	369	33.57	102	27.64	5.15	10.30	14.63	29.26	70.73
VanRaan	125	53.46	41	32.80	8.80	17.60	36.80	73.60	26.40
Vinkler	55	21.64	20	36.36	12.73	25.45	36.36	72.72	22.27
Waltmann	79	85.68	20	25.32	11.39	22.78	44.30	88.60	11.39
White	67	46.48	10	14.93	11.94	23.88	29.85	59.70	40.30
Zitt	39	31.62	12	30.77	15.38	30.77	48.72	97.44	2.56
Mean	135.9	42.92	36.0	26.56	10.50	20.99	32.83	64.13	38.36
SD	105.5	21.38	28.9	5.78	3.73	7.46	14.07	25.56	23.81

TP: total number of papers published in 1976–2021

Mean TC/TP: mean number of citations received to TP papers in 1976–2021

ECP, ECP%: number and per cent of papers, respectively cited more frequently than the mean, TC/TP

P(π)%, P(2π)%, P(h)%, P(2h)%, [(TP-P(2h)]%: per cent of papers in the π-set ($P(\pi) = \sqrt{TP}$), 2π-set, P(h)-set, P(2h)-set, and TP-P(2h)-set, respectively. The number of papers in the P(h)-set corresponds to the h-index

P(2h) + [TP-P(2h)] = TP, accordingly: P(2h)% + [TP-P(2h)]% = 100%

[a]There are less number of papers in the total set than 2h

[b]The data refer to: [TP–P(h)] number of papers

The percentage share of citations in the π-set takes 54.90% (SD = 11.36), in average, whereas that in the h-set is as high as 82.05% (SD = 11.58). The h-set contains 1.53 (SD = 0.23) times more citations than the π-core, in average. The number of citations in 2h-sets approaches to the total (94.25%, SD = 7.42). This fact may be attributed to ranking the papers by the number of citations and to the skewed distribution of citations over papers. The share of papers ranked higher than 2h, i.e. those in the (TP-2h) set, is equal to 38.36% (SD = 23.81) (Table 3), whereas the share of citations to those papers is only 6.19% (SD = 7.18) (Table 4).

The data in Table 5 reveal significant positive correlation ($r = 0.76$) between TP and TC. The share of citations in the elite subsets, C(π)%, C(2π)%, C(h)%, and C(2h)% was found to *correlate negatively* (and mostly significantly) both with total papers, TP and citations,

Table 4 Total number of citations (TC) obtained to TP papers, and percentage share of citations (C) received by papers in π (C(π)%), 2π (C(2π)%), h (C(h)%), 2h-set (C(2h)%), and that in the set of total papers (TP) minus papers in the 2h-set ([TC-C(2h)]%)

Name	TC	C(π)%	C(2π)%	C(h)%	C(2h)%	[TC-C(2h)]%	C(h)%/ C(π)%
Braun	2879	56.44	71.45	81.45	93.19	6.81	1.44
Cronin	4214	51.66	67.37	65.95	84.55	15.45	1.28
Egghe	4129	59.80	73.43	74.17	82.51	17.49	1.24
Glänzel	9241	38.87	53.84	69.13	90.07	9.93	1.78
Ingwersen	2629	67.21	79.19	87.45	98.67	1.33	1.30
Leydesdorff	16,051	43.64	56.43	70.11	88.18	11.82	1.61
Martin	1390	57.12	80.07	95.47	100.00	4.53[a]	1.67
McCain	2828	76.38	88.43	93.53	99.79	0.21	1.22
Moed	5875	48.03	66.49	84.31	99.86	0.14	1.76
Narin	5194	60.80	81.02	95.80	100.00	4.20[a]	1.58
Person	2140	67.80	83.46	92.24	99.91	0.09	1.36
Rousseau	5247	47.34	60.93	61.62	80.08	19.94	1.30
Schubert	4624	48.29	66.61	78.16	95.74	4.26	1.62
Small	2987	51.09	74.39	89.92	99.80	0.20	1.76
Thelwall	12,387	38.30	53.39	61.61	78.93	21.07	1.61
VanRaan	6683	44.82	61.89	84.06	99.28	0.72	1.88
Vinkler	1190	44.12	64.96	78.40	96.39	3.61	1.78
Waltmann	6737	64.54	81.67	95.74	100.00	0.00	1.48
White	3114	76.40	86.93	90.08	98.00	2.00	1.15
Zitt	1233	55.39	77.94	91.81	100.00	0.00	1.78
Mean	5038.6	54.90	71.50	82.05	94.25	6.19	1.53
SD	3806.1	11.36	10.83	11.58	7.42	7.18	0.23

The last column gives the rate of citations in h-set related to that in π-set

[a]As TP < P(2h), the value refers to: TC – C(h)

Table 5 Spearman rank-order correlation coefficients between the total number of papers, TP and total number of citations, TC = 100% of the whole publication set of the scientists studied and percentage share of citations to papers in the π-set, C(π)%, 2π-set, C(2π)%, h-set, C(h)%, and 2h-set, C(2h)%

	TP	TC	C(π)%	C(2π)%	C(h)%	C(2h)%
TP	1.00					
TC	0.76*	1.00				
C(π)%	– 0.57*	– 0.51*	1.00			
C(2π)%	– 0.73*	– 0.56*	0.96*	1.00		
C(h)%	– 0.86*	– 0.44	0.73*	0.85*	1.00	
C(2h)%	– 0.85*	– 0.39	0.50*	0.67*	0.93*	1.00

The correlation coefficients with „*" are significant at $p < 0.05$

TC. The correlation coefficient is however not significant for the TC-C(h)% (r = − 0.44) and TC-C(2h)% relation (r = − 0.39). The share of citations in the elite sets, e.g. C(π)% and C(h)% correlates with each other positively and significantly (r = 0.73).

Difference in citation rate (C/P) of papers in elite sets

In agreement with the expectations, the citation rate of papers, (C/P) in the individual elite sets decreases as follows: π-set: 232.13 (SD = 107.31), 2π-set: 151.48 (SD = 68.50), h-set: 111.57 (SD = 37.36), 2h-set: 64.20 (SD = 20.83) (Table 6). The difference between the mean of the total set (M = TC/TP = 42.92, SD = 21.38) and that of the elite sets, C/P(π), C/P(2π), C/P(h), C/P(2h) was found highly significant ($p < 0.003$). Similarly, all differences between the mean C/P value of the elite sets are highly significant ($p < 0.008$).

The Spearman correlation coefficients between C/P(π), C/P(2π), C/P(h), C/P(2h) of the elite sets are positive and significant (Table 7).The C/P values referring to the

Table 6 Mean citations per paper values, M = (C/P) for papers in the total set (TP) of the studied scientists, C/P values in the elite subsets, and Single plus First Authorship Rates (SFAR%)

Name	M = TC/TP	C/P					
		π-set	2π-set	h-set	2h-set	[TP-P(2h)]-set	SFAR%
Braun	35.99	180.56	114.28	83.75	47.91	8.17	78.75
Cronin	15.16	128.06	83.50	86.84	55.67	3.04	88.13
Egghe	21.23	166.86	102.43	99.90	55.57	5.42	99.46
Glänzel	36.90	224.50	155.47	116.15	75.66	6.51	41.04
Ingwersen	32.06	196.33	115.67	88.42	49.88	1.17	59.76
Leydesdorff	50.96	389.22	251.61	175.83	110.57	10.15	57.12
Martin	51.48	158.80	111.30	82.94	–	5.73[a]	55.56
McCain	43.51	270.00	156.31	120,23	64.14	0.29	73.85
Moed	52.93	255.55	177.55	117.93	69.85	0.30	40.54
Narin	103.08	451.14	300.57	177.71	–	9.90[a]	48.00
Person	46.52	207.29	127.57	94.00	50.90	0.50	34.78
Rousseau	18.67	146.17	94.03	89.81	60.30	4.96	25.62
Schubert	30.83	186.08	128.33	100.39	61.49	2.53	52.67
Small	46.64	190.75	138.88	103.31	57.33	0.50	82.81
Thelwall	33.57	249.68	174.03	141.33	90.53	10.00	51.49
VanRaan	53.46	272.27	188.00	122.13	72.12	1.45	41.00
Vinkler	21.64	75.00	55.21	46.65	28.68	2.87	100.00
Waltmann	85.68	483.11	305.67	184.29	96.24	2.88	40.51
White	46.48	297.38	169.19	140.25	76.30	2.30	85.07
Zitt	31.62	113.83	80.08	59.58	32.45	3.00	69.23
Mean	42.92	232.13	151.48	111.57	64.20	4.08	61.27
SD	21.38	107.31	68.50	37.36	20.83	3.36	22.11

[a]The value refers to the citation rate, C/P of papers in the [TP–P(h)] set

SFAR%: Single plus First Authorship Rate i.e., per cent of papers within the total (TP) where the studied scientist is the first or only author

– There are no data

Table 7 Spearman rank-order correlation coefficients between citation rate (C/P) values of the different elite sets and that of the total set (TC/TP)

	TC/TP	C/P			
		π-set	2*π*-set	*h*-set	2*h*-set
TC/TP	1.00				
C/P(*π*)-set	0.77*	1.00			
C/P(2*π*)-set	0.80*	0.97*	1.00		
C/P(*h*)-set	0.62*	0.93*	0.94*	1.00	
C/P(2*h*)-set	0.57*	0.84*	0.88*	0.95*	1.00

The correlation coefficients labelled with „*" are significant at $p < 0.05$

TP-P(2*h*)-set do not correlate with other C/P-data. This may be due to the relatively low and changing number of citations to papers in this set.

Figure 1 represents linear relationship between mean C/P of papers in the elite sets and mean citation rate (M = TC/TP) of the corresponding total sets. The slope of the lines increases in the rank: TP-P(2*h*), 2*h*, *h*, 2*π*, and *π*-set (0.026, 0.734, 1.296, 2.821, and 4.315, respectively). The citation rate of papers (C/P) in the Hirsch-core, termed here as C/P(*h*), corresponds to the A index suggested by Jin et al. (2007).

The mean share of single authored papers plus that signed by the studied scientist as first author (SFAR%) was found 61.27% (SD = 22.11) (Table 6). The dynamic range: 25.62% (Rousseau) – i.e. 74.38% of the papers are published in cooperation where the mentioned scientist is not the first author – 100.00% (Vinkler) – i.e. all papers are single or first authored.

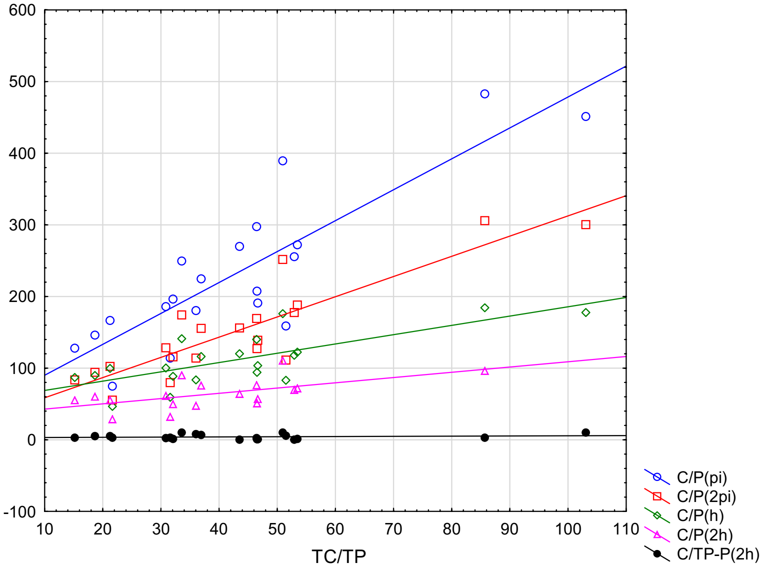

Fig. 1 Relationship between the citation rate of papers in the total set (TC/TP) against the citation (C) per paper (P) index of papers in the elite sets (*π* = *pi*, 2*π*, *h*, and 2*h*) and that in the [TP-P(2*h*)]-set

Study on the relationship between π and h-index and Mean Authorship Rate (MAR)

The Mean Authorship Rate (MAR) was calculated as a specific index according to the formula: MAR = TA/P, where TA is the total number of *authorships* of papers in the total set or in the corresponding elite sets (π, 2π, h, $2h$), further in (TP-2h)-set, and P is the number of papers in the corresponding set. The number of authorships of a *paper* is equal to the number of authors listed in the by-line. The number of authorships in a *set* can be obtained as the sum of authorships of the corresponding papers.

The following Mean Authorship Rate (MAR) indices were found for the total set of the scientists: MAR = 2.23 (SD = 0.52), whereas that for π, 2π, h and $2h$-set: 2.31 (SD = 0.59), 2.31 (SD = 0.64), 2.26 (SD = 0.53), 2.28 (SD = 0.55), respectively (Table 8). The MAR index for the papers cited rarely, i.e. that in (TP-2h)-set: 2.18 (SD = 0.73) (Table 8). Significance of the difference between Mean Authorship Rate (MAR) values referring to the different sets was also calculated. *No significant difference* (at $p < 0.05$) was found between the corresponding MAR values. Accordingly, both the most cited papers in the π-set and the less cited ones in the (TP-2h)-set show *similar number of authors*.

Table 8 Mean Authorship Rate (MAR) of papers in the total set, TP and in the elite sets: π-set, 2π-set, h-set, $2h$-set, and (TP-2h)-set

Name	Total set, TP		π-set		2π-set		h-set		$2h$-set		(TP-2h)-set	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Braun	2.64	0.99	2.67	0.50	2.94	1.11	3.00	0.98	2.88	0.95	2.08	0.88
Cronin	1.60	1.33	2.35	1.22	2.35	1.32	2.38	1.36	2.17	1.22	1.43	1.32
Egghe	1.42	0.63	1.57	0.65	1.57	0.63	1.55	0.63	1.55	0.68	1.36	0.60
Glänzel	2.70	1.20	2.25	0.77	2.38	0.87	2.44	0.84	2.55	1.01	2.80	1.31
Ingwersen	2.64	2.44	1.67	0.50	3.22	3.29	2.69	2.85	2.67	2.28	2.54	2.74
Leydesd	2.29	1.14	2.22	1.06	2.03	0.89	2.08	0.85	2.17	0.89	2.37	1.29
Martin	2.30	1.77	1.80	1.10	1.70	0.95	1.88	1.20	–	–	2.91*	2.30
McCain	1.65	0.96	1.88	0.99	1.63	0.81	1.77	1.07	1.57	0.87	1.81	1.12
Moed	2.77	1.53	2.45	1.29	2.50	1.30	2.48	1.33	2.73	1.40	2.89	1.91
Narin	2.42	0.99	3.14	0.69	2.64	0.84	2.54	0.88	–	–	2.27*	1.12
Person	2.39	1.77	2.43	0.98	2.00	0.88	2.24	0.89	2.50	1.81	1.25	0.50
Rousseau	2.32	1.15	2.53	1.37	2.32	1.15	2.31	1.13	2.19	1.01	2.37	1.20
Schubert	2.25	1.70	2.33	0.65	3.25	3.44	3.00	2.88	2.86	2.09	1.72	0.97
Small	2.16	2.31	1.88	0.83	1.75	0.78	1.69	0.97	1.79	1.21	3.75	4.52
Thelwall	2.31	1.27	2.78	1.26	2.49	1.19	2.33	1.05	2.25	1.05	1.29	1.36
VanRaan	2.38	1.42	3.27	2.80	3.09	2.22	2.78	1.82	2.50	1.54	2.15	1.03
Vinkler	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00	1.00	0.00
Waltmann	2.95	1.52	3.44	2.74	3.33	2.09	2.97	1.65	2.99	1.55	2.43	1.27
White	1.66	1.17	2.13	1.64	1.75	1.29	1.75	1.21	1.85	1.21	1.85	1.31
Zitt	2.85	1.87	2.37	0.52	2.25	0.75	2.37	0.83	2.84	1.90	3.30*	2.43
Mean	2.23		2.31		2.31		2.26		2.28		2.18	
SD	0.52		0.59		0.64		0.53		0.55		0.73	

Mean Authorship Rate: MAR = TAP/P, where TAP is total authorship and P is the number of papers in the corresponding set

*The data refer to: [TP–P(h)] number of papers

Table 9 shows the Spearman rank-order correlation coefficients between the MAR index of the total set (TP) and that in the different elite sets with total number of papers (TP, Table 3) and citations (TC, Table 4). Relations of MAR index with *h* and π-index of the total set (Table 10, FA-set) are also given.

The data show that the correlation coefficients between the MAR index of the different sets are significant (at $p < 0.05$). The MAR index of total papers (TP in FA-set) however does not correlate significantly with total number of papers (TP) ($r = -0.11$). The MAR index does not correlate significantly neither with total citations (TC) ($r = 0.31$) nor with π-index ($r = 0.23$) and *h*-index ($r = 0.20$) at $p < 0.05$. Accordingly, the *authorship rate certainly has no influence on π-index and h-index of total set.*

It seems to be interesting that the mean authorship rate (MAR) referring to π and 2π-set is relatively weekly but *significantly related* with both TC ($r = 0.56$, 0.52, resp.), π-index (0.51, 0.45, resp.) and *h*-index (0.45, 0.52, resp.) of the total set (TP). At the same time the MAR index (referring to *h*-set and 2*h*-set) show no significant correlation with *h*-index and π-index of the total set (TP). Nevertheless, the difference between the significant and non-significant correlation coefficients is low.

The Spearman rank-order correlation coefficient of π-index and *h*-index with total papers (TP) ($r = 0.68$, 0.85, resp.) and total citations (TC) ($r = 0.89$, 0.95, resp.) is relatively high and significant. This is because both impact indices depend on the *number of papers and citations, plus on the distribution of citations over papers* (Iglesias & Pecharromán, 2007).

Effect of the number and rank of authors on the h and π-index

The effect of the number and rank of authors on two elite set indicators (*h*-index, π-index) was studied according to Full Authorship Method (FAM), Partial Authorship Method (PAM), Practical Rank Score Method (PRSM), and Single plus First Authorship Method (SFAM) (see Tables 2 and 10). The set of papers obtained by the mentioned methods is termed as FA, PA, PRS, and SFA-set, respectively.

Table 9 Spearman rank-order correlational coefficients between Mean Authorship Rate (MAR) in the different sets and total number of papers (TP), total number of citations (TC), π-index, *h*-index of the total set (for more explanation see the text)

	Mean Authorship Rate					TP	TC	π-index
	TP-set	π-set	2π-set	*h*-set	2*h*-set			
TP-set	1.00							
π-set	0.61*	1.00						
2π-set	0.66*	0.68*	1.00					
h-set	0.68*	0.67*	0.96*	1.00				
2*h*-set	0.86*	0.66*	0.86*	0.90*	1.00			
TP	− 0.11	0.18	0.29	0.19	− 0.01	1.00		
TC	0.31	0.56*	0.52*	0.42	0.26	0.74*	1.00	
π-index	0.23	0.51*	0.45*	0.31	0.15	0.68*	0.89*	1.00
h-index	0.20	0.45*	0.52*	0.43	0.24	0.85*	0.95*	0.84*

Correlation coefficients labelled with "*" are significant at $p < 0.05$

Table 10 π-index and h-index calculated for FA-set (i.e. the original total set, TP with full credit to the studied author), PA-set, PRS-set, and SFA-set of papers of the scientists studied

Name	FA-set (TP)				PA-set		PRS-set		SFA-set	
	π-index	h-index	2π-index	$2h_c$-index	π-index	h-index	π-index	h-index	π-index	h-index
Braun	16.25	28	20.57	26.83	6.45	16	6.99	18	9.66	24
Cronin	21.77	32	28.39	35.63	13.01	26	15.14	27	17.28	29
Egghe	23.36	29	28.68	32.23	19.44	25	20.63	27	23.36	29
Glänzel	35.92	55	49.75	82.23	20.59	30	21.38	35	22.59	40
Ingwersen	17.67	26	20.82	25.94	13.06	16	11.69	18	9.67	15
Leydesdorff	70.06	64	90.58	141.53	41.04	45	35.15	48	25.68	52
Martin	7.94	16	11.13	–	5.85	14	5.85	13	5.16	10
McCain	21.60	22	25.01	28.22	14.80	18	13.17	20	10.21	18
Moed	28.22	42	39.06	58.67	16.85	27	18.16	30	17.93	31
Narin	31.58	28	42.08	–	11.91	23	14.34	22	18.13	17
Person	14.51	21	17.86	21.38	7.59	15	7.35	14	6.67	8
Rousseau	24.84	35	31.97	42.02	11.66	23	10.04	22	4.44	18
Schubert	22.33	36	30.80	44.27	10.02	23	10.50	22	12.67	19
Small	15.26	26	22.22	29.81	11.77	24	12.33	25	13.76	25
Thelwall	47.44	54	66.13	97.77	20.12	36	24.85	38	33.97	38
VanRaan	29.95	46	41.36	66.35	17.86	31	20.43	28	17.65	24
Vinkler	5.25	20	7.73	11.47	5.25	20	5.25	20	5.25	20
Waltmann	43.48	35	55.02	67.37	20.13	23	19.20	24	17.27	20
White	23.79	20	27.07	30.52	14.27	17	17.06	18	23.79	19
Zitt	6.83	19	9.61	12.33	3.24	13	4.32	12	6.21	16
Mean	25.40	32.70	33.29	47.48	14.25	23.25	14.69	24.05	15.07	23.60
SD	15.30	13.39	20.29	33.38	8.24	7.97	7.67	8.83	8.12	10.59

Partial Authorship, (PA)-set: papers of which citations are distributed among the authors according to the number of authors, PRS-set: papers of which citations are distributed among the authors according to the Practical Rank Score Method (see Tables 1, 2), SFA-set: papers *authored* only by the studied scientist plus that of which *first author* is the scientist studied

$2h_c$: 0.01 times the number of citations obtained to papers in the set containing $2h$ number of papers

The mean π-index of the studied scientists taking into account their **FA-set** (i.e. total set where total credit of the papers is attributed to the scientist studied) was found as 25.40 (SD = 15.30) (Table 10). The dynamic range is rather wide: 5.25 (Vinkler) – 70.06 (Leydesdorff). The latter value is extremely high in scientometrics taking into account the relatively low number of journals and scientists working in this field. The h-index calculated for the mentioned scientist (h-index = 64) represents also an extremely high value. The mean h-index of the studied scientists was found as 32.70 (SD = 13.39). The dynamic range: 16–64.

The π, h, 2π, and $2h_c$-index of FA-set (the FA-set corresponds to TP-set, i.e. total set with attributing full credit to the researcher studied) depends linearly on total number of citations (TC) (Fig. 2). The slope of the lines increases in the following rank: h, π, 2π, and $2h_c$-index (0.0033, 0.0037, 0.0051, 0.0083, resp.). $2h_c$-index of a set of papers can be calculated as the number of citations received by $2h$ number of papers multiplied with 0.01. Accordingly e.g., if $h = 20$, and the first 40 papers received 1147 citations, $2h_c$-index = 11.47. The $2h_c$-set of the studied scientists contains 93.52% (SD = 7.79) of total citations, in average.

The **PA-set** of the scientists obtained by Partial Authorship Method, also contains all papers (TP) but the *rank* of the papers according to the number of citations *differs* from that in FA-set (see Table 2). This is because the number of citations obtained to a paper of the studied author was calculated as c_i/a_i, where c_i is the number of citations received and a_i is the *number of authors* of the i-th paper.

The mean π and h-index was found as 14.25 (SD = 8.24) and 23.25 (SD = 7.97), respectively. The dynamic range of π-index: 3.24 (Zitt) – 41.04 (Leydesdorff) and that for h-index: 13 (Zitt) – 45 (Leydesdorff) (Table 10).

The number of papers in the **PRS-set** is equal to that in the total set (TP in FA-set). Nevertheless, the PRS-method takes into account not only *the number but also the rank* of

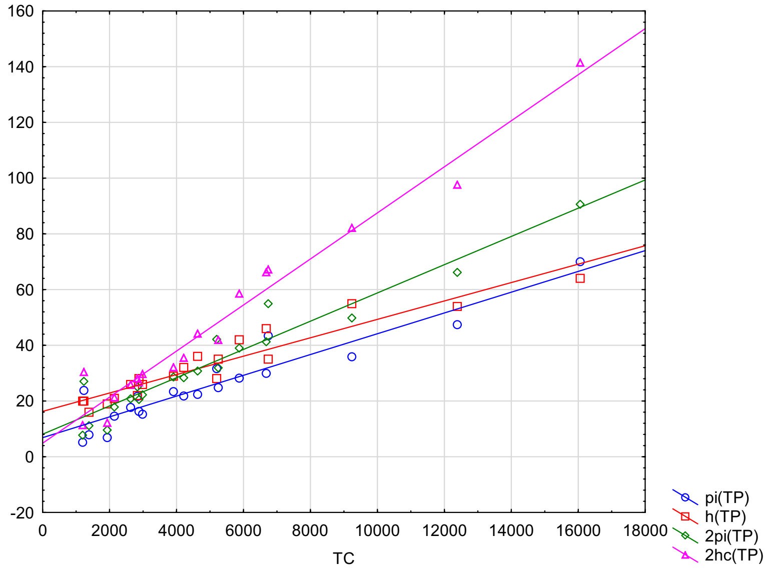

Fig. 2 Scatterplot of total citations (TC) against pi(TP) = π(TP), $2pi$(TP) = 2π(TP), h(TP), and $2h_c$(TP) index (number of papers in the $2h_c$-set = $2h$) calculated for the total set (TP = FA)

the assessed authors according to Tables 1 and 2. The mean π and h-index of the scientists studied was found as follows: 14.69 (SD = 7.67) and 24.05 (SD = 8.83), respectively. The dynamic range of π-index: 4.32 (Zitt) – 35.15 (Leydesdorff) and that for h-index: 12 (Zitt) – 48 (Leydesdorff) (Table 10).

According to the Single and First Authorship Method (SFAM) only those papers are taken into consideration which are signed by the studied scientist as *single author* or of which *first author* is the studied researcher (SFAR%, Table 6). Naturally, the number of those papers may be lower or significantly lower than the number of total publications (see Table 2). The mean percentage rate of **SFA** papers of the scientists studied is 61.27 (SD = 22.11) (Table 6). The mean π and h-index of the papers in the SFA-sets: 15.07 (SD = 8.12) and 23.60 (SD = 10.59), respectively (Table 10). The dynamic range of π-index: 4.44 (Rousseau) – 33.97 (Thelwall), and h-index: 8 (Person) – 52 (Leydesdorff).

From the data in Table 10 the corresponding significance values of the difference between the mean h and π-index of the sets obtained by Full Authorship Method (FAM), Partial Authorship Method (PAM), Practical Rank Score Method (PRSM), and Single plus First Authorship Method (SFAM) were calculated (Table 11).

The difference between the mean π-index of the sets obtained by Full Authorship Method (FAM), PAM, PRSM or SFAM was found *significant* ($p < 0.025$, bold values in the first column in Table 11). The conclusion is valid also for h-index (second column in Table 11). In contrast, *no significant difference* was found between the π and h-values obtained by PAM, PRSM and SFA-Method (values in italics and bold italics, resp., Table 11).

Comparing the h-index obtained by FAM with that obtained by PAM, PRSM, and SFAM, two trends may be observed (Table 10). There are persons with ***h*-index** of their papers in the **FA**-set that does not change at all or changes only *slightly* by applying any of the mentioned methods (PAM, PRSM, or SFAM): Cronin (**32**, 26, 27, and 29), Martin (**16**, 14, 13, and 10), McCain (**22**, 18, 20, and 18), Small (**26**, 24, 25, and 25), Vinkler (**20**, 20,

Table 11 Significance level of differences between the mean π and h-index of the different sets (see Table 10)

		FA-set		PA-set		PRS-set		SFA-set	
		π	h	π	h	π	h	π	h
FA-set	π	1.000							
	h	0.117	1.000						
PA-set	π	**0.007**	0.000	1.000					
	h	0.499	**0.007**	0.002	1.000				
PRS-set	π	**0.008**	0.000	*0.862*	0.002	1.000			
	h	0.734	**0.021**	0.001	***0.632***	0.001	1.000		
SFA-set	π	**0.011**	0.000	*0.753*	0.005	*0.880*	0.002	1.000	
	h	0.668	**0.022**	0.004	***0.778***	0.004	***0.885***	0.007	1.000

Significance levels in bold (e.g. **0.007**) refer to the difference between the mean π and h-index of sets obtained by Full Authorship Method (FA), Partial Authorship Method (PA), Practical Rank Score Method (PRS), and Single plus First Authorship Method (SFA)

Significance levels in italics (e.g. *0.862*) refer to the difference between the mean π-index of sets obtained by PA, PRS and SFA method

Significance levels in bold italics (e.g. ***0.632***) refer to the difference between the mean h-index of sets obtained by PA, PRS and SFA method

20, and 20), White (**20**, 17, 18, and 19). In contrast, there are scientists with significantly *higher* change in *h*-index: Glänzel (**55**, 30, 35, and 40), Ingwersen (**26**, 16, 18, and 15), Leydesdorff (**64**, 45, 48, and 52), Moed (**42**, 27, 30, and 31) Schubert (**36**, 23, 22, and 19), Waltman (**35**, 23, 24, and 20).

It may be assumed that π and *h*-index of the FA-set compared to that of PA, PRS, and SFA-set could be positively and significantly related to the Single plus First Authorship Rate (SFAR%) of the scientists. Therefore, the corresponding Spearman rank-order correlations were calculated (Table 12). In contrast to the expectations, for the FA, PA and PRS-set no significant and positive correlations were found. The correlation coefficient between SFAR% and SFA(π), SFA(h) is positive but not significant.

Figure 3 demonstrates the linear relationship of h(FA) with h(SFA), h(PRS), h(PA), and $2h$(FA). The slope of the lines: 0.685, 0.613, 0.562, and 2.358, resp.

Figure 4 demonstrates also linear relationship of π(FA) with 2π(FA), π(PA), π(PRS), and π(SFA). The slope of the lines increases by the following rank: π(SFA): 0.398, π(PRS): 0.463, π(PA): 0.500, and 2π(FA): 1.319.

From the above we may conclude that applying any of the part-impact methods, i.e. distributing credit of publications and citations among the coauthors, may decrease the value of *h and* π-index by *about of 55–60 and 70–75 percent,* respectively. However, the decrease is not homogeneous. Table 13 shows *the effect of decrease on the rank* of the scientists. The rank value of the scientists according to *h* or π-index (or both) obtained from the data of the (total) FA-set differs from that calculated for PA-set in several cases (e.g. Braun, Egghe, McCain, Narin, Rousseau, Small, Vinkler, and White). In contrast, there are scientists (e.g. Leydesdorff, Martin, Moed, Person, vanRaan, Waltmann, and Zitt) showing similar rank values. Similar differences may be observed in the rank comparing data from FA and PRS-set.

The discrepancies between the ranks by *h* or π-index obtained from the FA, PA, PRS and SFA-set of the authors may be demonstrated by the *mean difference* in the rank number of the scientists (Table 14, in brackets). As another characteristic index: the number of cases with a difference of *three or greater* in the rank number may be also applied (Table 14).

The rank by the indices obtained from FA-set against that calculated from PA-set changes at a greater extent for the scientists publishing together with relatively many

Table 12 Spearman rank-order correlation coefficients between the Single plus First Authorship Rate (SFAR%) and π-index, *h*-index of sets obtained by Full Authorship Method (FAM), Partial Authorship Method (PAM), Practical Rank Score Method (PRSM), and Single plus First Authorship Method (SFAM)

	SFAR%	π(FA)	h(FA)	π(PA)	h(PA)	π(PRS)	h(PRS)	π(SFA)
SFAR%	1.00							
π(FA)	– 0.46	1.00						
h(FA)	– 0.43	0.84*	1.00					
π(PA)	– 0.22	0.86*	0.73*	1.00				
h(PA)	– 0.28	0.74*	0.84*	0.64*	1.00			
π(PRS)	– 0.16	0.87*	0.76*	0.97*	0.71*	1.00		
h(PRS)	– 0.17	0.78*	0.89*	0.79*	0.89*	0.86*	1.00	
π(SFA)	0.05	0.76*	0.60*	0.81*	0.59*	0.91*	0.73*	1.00
h(SFA)	0.14	0.58*	0.74*	0.66*	0.75*	0.74*	0.88*	0.69*

*Significant at $p < 0.05$

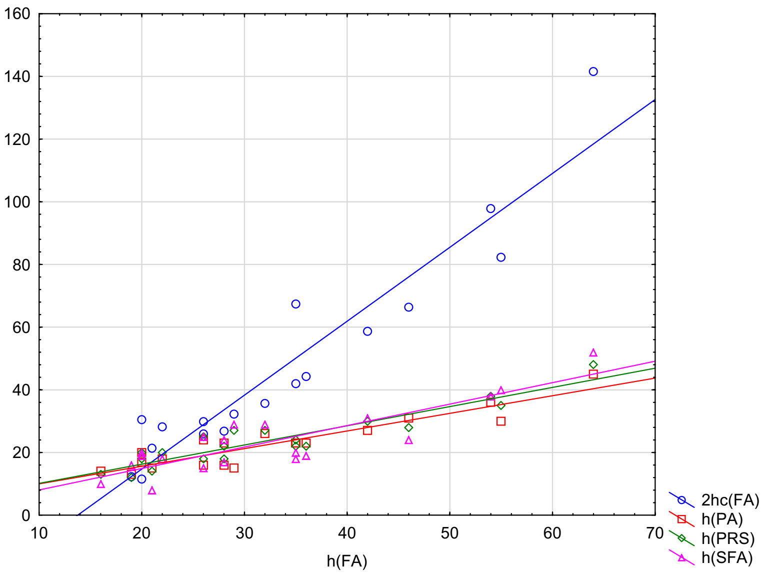

Fig. 3 Scatterplot of *h*(FA) index against $2h_c$(FA), *h*(PA), *h*(PRS), and *h*(SFA)-index of the studied scientists. (For FA, PA, PRS, and SFA see the text)

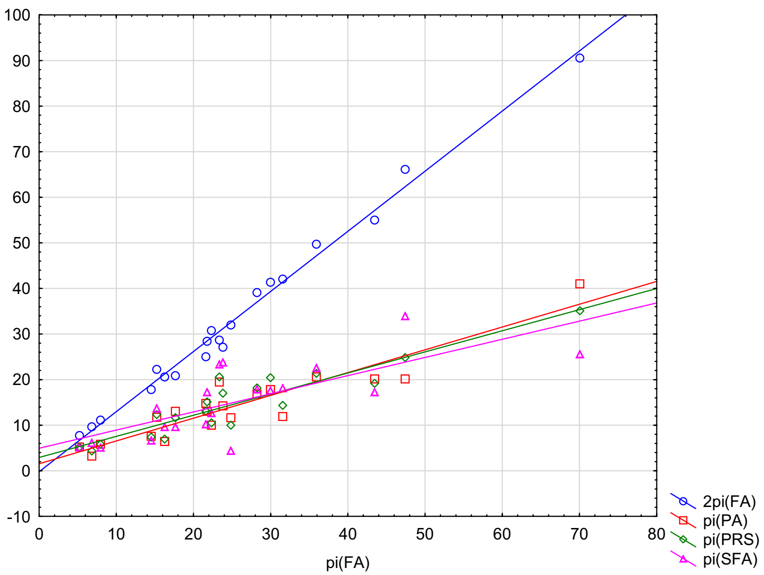

Fig. 4 Scatterplot of pi = π-values calculated with attributing full credit of papers (FA) to the author studied against 2pi(FA) = 2π(FA)-index and pi(PA) = π(PA)-index. The latter mentioned index is calculated with c_i/a_i part of citations (c_i) attributed to the studied scientist, where a_i is the number of authors of i-th paper. The pi(PRS) = π(PRS)-index refers to papers calculated with citations taking into account number and rank of authors, whereas the pi(SFA) = π(SFA) index is calculated according to the single and first authorship method

Table 13 Rank of the studied scientists according to π-index and h-index calculated from the data of FA-sets (i.e. all papers, TP attributing full credit to the studied author), PA-sets, PRS-sets, and SFA-sets (attributing only the relevant part of the credit to the author studied)

Name	FA-set (TP)		PA-set		PRS-set		SFA-set	
	π-index	h-index	π-index	h-index	π-index	h-index	π-index	h-index
Braun	15	11.5	17	14.5	17	16	15	8.5
Cronin	12	9	11	6	9	6.5	9	5.5
Egghe	10	10	5	17.5	4	6.5	4	5.5
Glänzel	4	2	2	4	3	3	5	2
Ingwersen	14	13.5	10	14.5	13	16	14	18
Leydesdorff	1	1	1	1	1	1	2	1
Martin	18	20	18	19	18	19	19	19
McCain	13	15	8	12	11	13.5	13	14.5
Moed	7	5	7	5	7	4	7	4
Narin	5	11.5	12	8.5	10	11	6	16
Person	17	16	16	17.5	16	18	16	20
Rousseau	8	7.5	14	8.5	15	11	20	14.5
Schubert	11	6	15	8.5	14	11	12	12.5
Small	16	13.5	13	7	12	8	11	7
Thelwall	2	3	4	2	2	2	1	3
VanRaan	6	4	6	3	5	5	8	8.5
Vinkler	20	17.5	19	11	19	13.5	18	10.5
Waltmann	3	7.5	3	8.5	6	9	10	10.5
White	9	17.5	9	13	8	16	3	12.5
Zitt	19	19	20	20	20	20	17	17

Partial Authorship (PA)-set: papers of which citations are distributed according to the number of co-authors; PRS-set: papers of which citations are distributed according to the Practical Rank Score Method (see Tables 1, 2); SFA-set: papers *authored* only by the studied scientist plus that of which *first author* is the scientist studied

coauthors. Applying the PRS-set, the change may be high for scientists ranked as third, fourth, fifth, etc. in the list of authors in the byline of the paper.

The SFA-set shows the greatest *mean differences in rank numbers* and number of *differences of three or higher than three* of h-index (3.37, 13, respectively). There are scientists with extreme great difference in h or π-index (or both), e.g. Egghe (FA: 10, 10, SFA: 4, 5.5,

Table 14 Number of cases with a difference of three or greater in the rank numbers of the studied scientists by h-index and π-index obtained from data of FA, PA, PRS, and SFA-set. Mean change in the rank number (in brackets) of the scientists by changing the set from which the impact indices are calculated

Set of papers	FA		PA		PRS	
	h	π	h	π	h	π
PA	8 (2.87)	7 (2.20)				
PRS	6 (2.17)	7 (2.10)	2 ((1.77)	3 (1.10)		
SFA	13 (3.37)	7 (2.65)	8 (3.05)	9 (2.85)	7 (2.10)	6 (1.90)

resp.), Rousseau (FA: 8, 7.5, SFA: 20, 14.5, resp.), White (FA: 9, 17.5, SFA: 3, 12.5, resp.), whereas the rank numbers of other persons (e.g. Leydesdorff, Martin, and Moed) are practically unchanged. The reason may be that the number of journal papers published solely, without coauthors may strongly depend on the personality and possibilities of the scientist. Similarly, the rank of authors in the by-line of journal papers may depend on several non-scientific factors. There are senior scientists who insist on the first position whilest others prefer the last one. The measure of contribution to the paper that is acknowledged by authorship differs from laboratory to laboratory (Vinkler, 1993).

In applying PA and PRS-set for calculating h-index or π-index, low number of differences of three or greater between the corresponding rank values: 2 and 3, respectively, and low mean rank differences: 1.77 and 1.10, respectively can be observed (Table 14). Accordingly, it may be concluded that the difference between the the respective indices is relatively small in most cases (see Table 10).

Summary and conclusions

In assessing publication performance of individuals, the importance of dividing the total set of journal papers of the studied scientists according to science fields was demonstrated in an earlier publication (Vinkler, 2021). The mentioned selection seems to be important because the *bibliometric features* (Vinkler, 1988) depend on the field. It is known that both quantity (e.g. number of papers) and impact indicators (e.g. number of total citations, citations/paper, Hirsch-index, π-index, etc.) strongly depend on the field. Without separating the journal papers of similar bibliometric features, the evaluation of total sets would yield incorrect results.

Scientific investigation is conducted in natural sciences mostly as teamwork. Consequently, multiauthored publications are common phenomenon. However, there are many situations in the practice (e.g. selection of applicants for granting, recommendation of persons for job, rewarding scientists with prices, study of performance of a country or laboratory, etc.) when *individual performance* should be determined. Consequently, it is relevant to distribute the credit of journal publications and citations among the coauthors.

For evaluation of individuals, teams or countries impact indicators of the *elite sets* of publications are preferably applied, recently. In the present study the following elite sets are applied: $\sqrt{P}$, h, $2\sqrt{P}$, and $2h$ number of publications ranked by decreasing citation frequency of the journal papers of 20 scientometricians (P is equal to the total number of papers and h is equal to h-index).

The data in Table 3 show that the mean *percentage rate of papers* in the elite sets *increases* according to the rank: P(π), P(2π), ECP, P(h), P($2h$), where P(π) is the number of papers in $\sqrt{P}=\pi$-set, and P(h) is that in h-set. The papers in ECP-set are cited more frequently than the mean TC/TP value of the corresponding total set. The mean share of papers in h-set (32.83%, SD = 14.07) was found 3.13 times higher than that in π-set (10.50%, SD = 3.73). The difference between the mean P(π)% and P(h)% values is *highly significant* ($p < 0.001$).

The share of citations in the elite subsets, C(π)%, C(2π)%, C(h)%, and C($2h$)% was found to *correlate negatively* with total papers, TP and total citations, TC (Table 5). C(π)% and C(h)% correlates with each other positively and significantly ($r = 0.73$).

In agreement with the expectations, it was found that the mean citation rate of papers (C/P) in the individual elite sets decreased as follows: π-set: 232.13 (SD = 107.31), 2π-set: 151.48 (SD = 68.50), h-set: 111.57 (SD = 37.36), $2h$-set: 64.20 (SD = 20.83).

The Mean Authorship Rate (MAR) index was calculated for the total and for the elite sets of papers of the studied scientists. The calculations show that MAR of total sets does not correlate significantly with total number of papers (TP) ($r = -0.11$), total citations (TC) ($r = 0.23$), π-index ($r = 0.23$), and h-index ($r = 0.20$) (Table 9). Accordingly, mean *authorship rate of papers in the total set may not influence the mentioned indices.* Nevertheless, the low but significant correlation coefficients between the number of authors of papers in the π-set and π-index, h-index ($r = 0.51$, 0.45, respectively) may indicate some impact of the number of authors. In contrast, no significant correlation was obtained between the number of authors in h and $2h$-set with π-index ($r = 0.31$, 0.15, respectively). Similarly, no significant correlation was detected between MAR in h or $2h$-set and h-index ($r = 0.43$, 0.24, respectively).

According to the results presented here, in analysing publication impact of scientists, teams or countries distributing citations among the coauthors is highly relevant. A recent paper of Ioannidis et al. (2019) also points to the importance of distributing credit among coauthors. The study of Egghe (2008) gives a mathematical theory of h and g-index in case of fractional counting of authorship.

Scientometric evaluations can be performed by taking into account several viewpoints. The type of the part-impact method applied, should be selected accordingly. If the evaluation emphasizes e.g., the importance of *cooperation* between scientists or teams, we have to select the total set (FA). This way evaluating e.g. grant applications, data obtained from the FA-set should be preferably applied. If we were interested in the personal contribution by a scientist, PA, PRS or SFA-set should be used. The contribution of the person to be assessed is emphasized by using PRS-set, assuming that rank of coauthors runs parallel with measure of contribution. The greatest stress is laid on personal contribution by applying the SFA-method. Anyway, it would seem most appropriate to calculate the elite set indicators according to each method. Possible weightings may be made according to relevant expert decisions.

Several important requirements for the evaluation processes (e.g. specifying the purpose of the assessment, specifying the system to be assessed, setting criteria, selecting methods, specifying indicators for each criterion, setting the time-period, selecting appropriate standards, etc. (Moravcsik, 1988; Vinkler, 2010a) cannot be tackled here.

Funding Open access funding provided by ELKH Research Centre for Natural Sciences.

References

Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation, 12*, 159–170.

Bornmann, L., Leydesdorff, L., & Mutz, R. (2013). The use of percentiles and percentile rank classes in the analysis of bibliometric data: Opportunities and limits. *Journal of Informetrics, 7*, 158–165.

Bornmann, L., & Marx, W. (2014). How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations. *Scientometrics, 98*, 487–509.

Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H.-D. (2011). A multilevel meta-analysis of studies reporting correlations between the *h* index and 37 different *h* index variants. *Journal of Informetrics, 5*, 346–359.

Egghe, L. (2006). Theory and practice of the *g*-index. *Scientometrics, 69*, 131–152.

Egghe, L. (2007). Dynamic *h*-index: The Hirsch index in function of time. *Journal of the American Society for Information Science and Technology, 58*, 452–454.
Egghe, L. (2008). Mathematical theory of the *h*- and *g*-index in case of fractional counting of authorship. *Journal of the American Society for Information Science and Technology, 59*(10), 1608–1616. https://doi.org/10.1002/asi.20845
Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 16569–16572.
Iglesias, J., & E. & Pecharromán, C. (2007). Scaling the h-index for different scientific ISI fields. *Scientometrics, 73*(3), 73–320.
Ioannidis, J. P. A., Baas, J., Klavans, R., & Boyack, K. W. (2019). A standardized citation metrics author database annotated for scientific field. *Plos Biology, 17*(8), e3000384. https://doi.org/10.1371/journal.pbio.3000384
Jin, B., Liang, L., Rousseau, R., & Egghe, L. (2007). The *R*- and *AR*-indices: Complementing the *h*-index. *Chinese Science Bulletin, 52*, 855–863. https://doi.org/10.1007/s11434-007-0145-9
Kolthun, V. & Hafner, D. (2021). The *h*-index is no longer an effective correlate of scientific reputation. *PLoS ONE*, June 28.
Leydesdorff, L. (2012). Alternatives to the journal impact factor: I3 and top-10% (or top-25%?) of the most highly cited papers. *Scientometrics, 92*, 355–365.
Lukovits, I., & Vinkler, P. (1995). Correct credit distribution: A model for sharing credit among coauthors. *Social Indicators Research, 36*, 91–98.
Moravcsik, M. J. (1988). Some contextual problems of science indicators. In A. F. J. Van Raan (Ed.), *Handbook of quantitative studies of science and technology* (pp. 11–30). Elsevier Science Publishers B. V.
Oberesch, E. & Groppe, S. (2017). The m_f index: A citation-based multiple factor index to evaluate and compare the output of scientists. *Open Journal of Web Technologies (OJWT) 4(1).* http://www.ronpub.com/ojwt
Osório, A. (2018). On the impossibility of a perfect counting method to allocate the credits of multi-authored publications. *Scientometrics, 116*, 2161–2173.
Plomp, R. (1990). The significance of the number of highly cited papers as an indicator of scientific prolificacy. *Scientometrics, 19*, 185–197.
Prathap, G. (2021). Letter to the editor: Dimensionless citation indicators for fractional counting. *Scientometrics, 126*, 8765–8769.
De Solla Price, D. J., & Beaver, D. (1966). Collaboration in an invisible college. *American Psychologist, 21*, 1011–1018.
Sangwal, K. (2022). Comparative study of scaling parameters and research output of selected highly- and moderately-cited individual authors. *Journal Computer Sciences Institute, 23*, 152–164.
Schreiber, M. (2010). Twenty Hirsch index variants and other indicators giving more or less preference to highly cited papers. *Annalen Der Physik (berlin), 52*, 536–554.
Schreiber, M., Malesios, C. C., & Psarakis, S. (2012). Exploratory factor analysis for the Hirsch index, 17 h-type variants, and some traditional bibliometric indicators. *Journal of Informetrics, 6*, 347–358.
Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science, 43*, 628–638.
Todeschini, R., & Baccini, A. (2016). *Handbook of bibliometric indicators—Quantitative tool for studying and evaluating research* (pp. 1–512). Wiley-VCH.
Vinkler, P. (1988). Bibliometric features of some scientific subfields and the scientometric consequences therefrom. *Scientometrics, 14*, 453–474.
Vinkler, P. (1993). Research contribution, authorship and team cooperativeness. *Scientometrics, 26*, 213–230.
Vinkler, P. (2009). The *π*-index. A new indicator for assessing scientific impact. *Journal of Information Science, 35*, 602–612.
Vinkler, P. (2010a). *The evaluation of research by scientometric indicators* (pp. 1–313). Chandos Publishing.
Vinkler, P. (2010b). The *πv*-index: A new indicator to characterize the impact of journals. *Scientometrics, 82*, 461–475.
Vinkler, P. (2017a). The size and impact of the elite set of publications in scientometric assessments. *Scientometrics, 110*, 163–177.
Vinkler, P. (2017b). Core indicators and professional recognition of scientometricians. *Journal of the Association for Information Science and Technology, 68*, 234–242.
Vinkler, P. (2021). Evaluation of publications by the part-set method. *Scientometrics, 126*, 2737–2757.
Wagner, C. S., Zhang, L., & Leydesdorff, L. (2022). A discussion of measuring the top-1% most-highly cited publications: quality and impact of Chinese papers *Scientometrics 127,* 1825–1839.