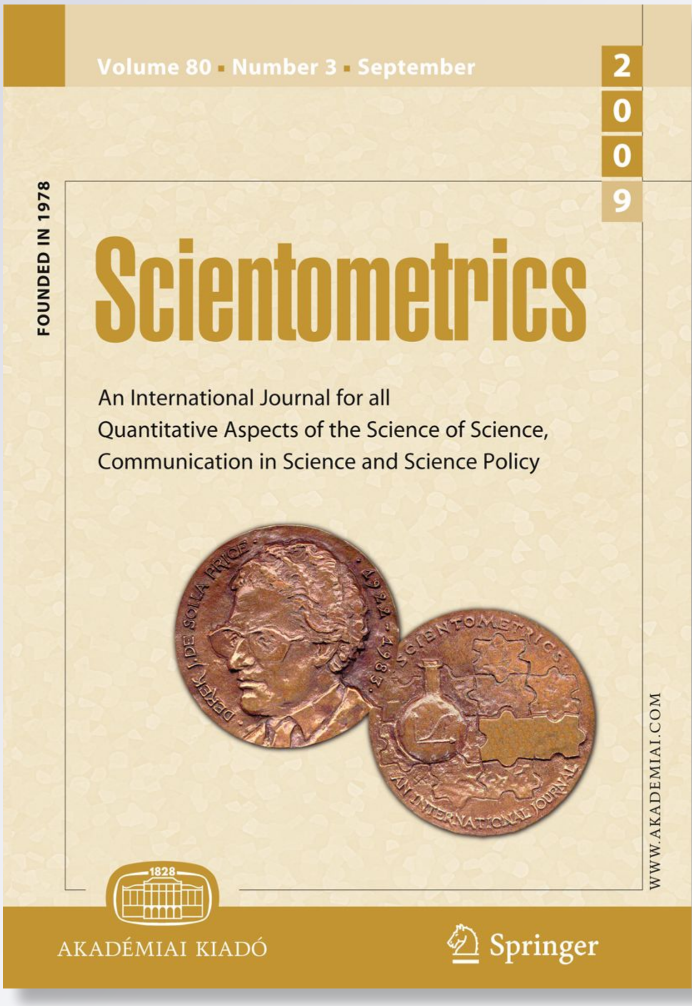
Research evaluation. Part I: productivity and citedness of a German medical research institution

A. Pudovkin, H. Kretschmer, J. Stegmann & E. Garfield

Scientometrics
An International Journal for all Quantitative Aspects of the Science of Science, Communication in Science and Science Policy

ISSN 0138-9130

Scientometrics
DOI 10.1007/s11192-012-0659-z



Scientometrics
DOI 10.1007/s11192-012-0659-z

Research evaluation. Part I: productivity and citedness of a German medical research institution

A. Pudovkin · H. Kretschmer · J. Stegmann · E. Garfield

Received: 23 January 2012

Abstract An evaluation exercise was performed involving 313 papers of research staff (66 persons) of the Deutsche Rheuma-Forschungszentrum (DRFZ) published in 2004–2008. The records and citations to them were retrieved from the *Web of Science* (Thomson Reuters) in March 2010. The authors compared productivity and citedness of "group leaders" vs. "regular scientists", of "male scientists" vs. "female scientists" using citation-based indexes. It was found that "group leaders" are more prolific and cited more often than "regular scientists", the same is true considering "male" vs. "female scientists". The greatest contrast is observed between "female leaders" and "female regular scientists". The above mentioned differences are significant in indexes related to the number of papers, while values of indexes characterizing the quality of papers (average citation rate per paper and similar indexes) are not substantially different among the groups compared. The mean value of percentile rank index for all the 313 papers is 58.5, which is significantly higher than the global mean value of about 50. This fact is evidence of a higher citation status, on average, of the publications from the DRFZ.

A. Pudovkin
Institute of Marine Biology, Far East Branch, Russian Academy of Sciences, Vladivostok 690041, Russia
e-mail: aipud@online.ru

H. Kretschmer (✉)
Faculty of Business Administration/Business Computing, University of Applied Sciences, Bahnhofstrasse, 15745 Wildau, Germany
e-mail: kretschmer.h@onlinehome.de

J. Stegmann
Berlin, Germany
e-mail: johannes.stegmann@alice.de

E. Garfield
Chairman Emeritus, Thomson Reuters Professional, Formerly Institute for Scientific Information®
(ISI®), 1500 Spring Garden Street, Philadelphia, PA 19130-4067, USA
e-mail: garfield@codex.cis.upenn.edu

Published online: 12 February 2012

Keywords Evaluation · Productivity · Citations · Gender · Group leaders · Percentile rank index

Mathematical Subject Classification 62 · 68 · 91 · 94

Introduction

Objective quantitative assessment of scientific productivity is an important though a highly controversial topic. Many indexes have been suggested for this purpose and critically discussed (Bar-Ilan 2008; Hirsch 2005; Ioanidis et al. 2007; Tijssen et al. 2009; Sanz-Casado et al. 2009; Wallin 2005). They are based on either the number of publications, or on the number of citations to these publications, or on indexes of journal impact or on a combination of them.

The authors performed an evaluation exercise of productivity and citation impact of the research staff of the Deutsche Rheuma-Forschungszentrum (DRFZ) in 2004–2008. A variety of citation indexes (Table 1) were used in an attempt to determine which are more informative in revealing the differences in performance of researchers who are "group leaders" and those in subordinate positions, who we designate as "regular scientists". Similarly, we were interested in gender differences.

Our study had two main goals:

(a) to find the characteristics which better differentiate researchers in two respects—leader vs. regular scientist and gender.

Table 1 Characters and their abbreviations

No	Abbreviation	Meaning
1	ASI99	Author superiority index* at 99th percentile
2	ASI95	Author superiority index*at 95th percentile
3	ASI75	Author superiority index*at 75th percentile
4	ASI50	Author superiority index*at 50th percentile
5	NumP	Number of papers (per author)
6	sumC	Sum of cites to papers (per author)
7	avC	Average number of citations (per paper, per author)
8	sumIF	Sum of IF values of journals where author's papers are published
9	avIF	Average journal IF (per paper, per author)
10	avPRI	Average percentile rank index* (per paper, per author)
11	*h*-Index	*h*-Index (after Hirsch 2005), per author
12	av%75	Average percentage of papers with PRI* $\geq$75 (per author)
13	avPRI*IF	Average product of PRI* by journal IF (per author)
14	sumC/y	Sum of age-corrected citations (per author)
15	avC/y	Average number of citations divided by the years after publication (per paper per author)
16	avNumAuth	Average number of authors (per paper, per author)
17	sumPRI	Sum of PRIs* of papers (per author)
18	sumPRI*IF	Sum of products of PRI* by journal IF (per author)

*see Appendix

(b) to see if the indexes—percentile rank index (PRI) and author superiority index (ASI)—suggested earlier by two of us (Pudovkin and Garfield 2009) have some advantage in revealing productive or potentially productive scientists compared to the commonly used characteristics of productivity and impact.

Data and methods

We created a data set of publications produced by the DRFZ in 2004–2008. We extracted from the *Web of Science* (Thomson Reuters [http://science.thomsonreuters.com/isi/]) all the publications by the staff of this institution and the numbers of citations to these publications. Data retrieval was performed in March 2010. Thus the citation data refer to this date. Consulting the web-site of this institution we identified the authors, their positions and gender. There are 313 papers in our data base, authored and/or co-authored by 66 DRFZ scientists. The sizes of the groups of researchers compared (leaders and regular scientists, male, and female) are shown in Table 2A, B. The 313 papers were published in 96 domestic and international journals, with impact factors (IF) ranging from 0.084 to 47.400, the median and quartiles being 4.226, 2.058, and 6.956. Citation frequencies range from 0 to 229, the median and quartiles being 9, 2, and 26.

The citation-based indexes we used are presented in Table 1. Some of these only characterize productivity of an author (NumP), some correlate with the quality of publications (avC, avIF, avPRI, avPRI*IF, and av%75), while others depend on both productivity and quality (*h*-index, sumIF, sumPRI, sumPRI*IF, ASI99, ASI95, ASI75, and ASI75).

The size of the differences between the compared groups were estimated by the difference index (DI): DI $= (x_1 - x_2)/\mathrm{SD}_{1\&2}$, x_1 and x_2 being the means in the compared groups 1 and 2, $\mathrm{SD}_{1\&2}$ being the averaged standard deviation. Statistical significance of the differences was estimated by the Student *t*-test.

Results and discussion

Table 2A gives the overall average values and the averages for the groups: "leaders" and "scientists", "male" and "female". One can see that, on average, the output of "leaders" is more substantial than that of the "regular scientists" both in number of papers and in citation impact. The same is true for the comparison of men vs. women. The mean DIs are 0.59 and 0.47.

Table 2B compares the sub-groups: "male leaders" vs. "female leaders", "male leaders" vs. "male scientists", etc. One can see that the largest average contrast (DI $=$ 0.87) is observed between "female leaders" and "female scientists", while the other contrasts are much smaller. Though the overall pattern of differences is consistent in sub-groups comparisons: leaders are more efficient than regular scientists, men are more efficient than women.

The overall significance of differences (leaders vs. scientists, men vs. women) is beyond doubt. The differences in individual parameters are less significant or non-significant, most probably because of the small sample sizes. It is not easy to perform a valid overall significance tests (over all the parameters) as many parameters are correlated.

Table 2 Bibliometric indexes for 66 scientists of the DRFZ

A[a]

Parameter	Lea	Sci	Difference index	M	F	Difference index	Mean	Me	Q1	Q3
Sample size	21	45	21 & 45	32	30	32 & 30	66	66	66	66
ASI99	0.60	0.17	0.60**	0.500	0.100	0.60**	0.30	0	0	0
ASI95	2.00	0.87	0.45*	1.781	0.700	0.48	1.21	0	0	1
ASI75	6.95	3.17	0.48*	6.281	2.467	0.51**	4.32	1	0	5
ASI50	10.85	4.30	0.71***	8.906	3.800	0.55**	6.29	3	1	7.5
NumP	18.50	6.30	0.98****	13.875	6.400	0.57**	10.00	5	2	11.5
sumC	432.6	142.1	0.82***	352.7	116.7	0.67***	230.1	69.5	19	216.5
avC	23.24	16.69	0.44*	19.97	17.20	0.19	18.67	16.30	7.63	24.65
sumIF	121.10	34.00	1.18****	89.61	32.80	0.73**	60.39	29.31	10.40	70.32
avIF	7.45	5.91	0.38	6.71	5.98	0.20	6.38	5.90	4.30	7.48
av%75	0.32	0.34	−0.07	0.36	0.31	0.16	0.33	0.29	0	0.50
h-Index	9.25	4.02	0.96****	7.438	3.900	0.61**	5.61	4	2	6
avPRI	59.2	53.94	0.22	58.56	51.57	0.32	55.53	56.45	45.00	70.43
avPRI*IF	390.9	301.8	0.47*	354.7	295.6	0.31	330.2	344.1	202.1	424.9
avC/y	4.49	5.90	−0.35	5.44	4.26	0.32	4.98	4.26	2.53	6.64
sumC/y	104.90	36.17	0.82***	87.14	28.10	0.71**	57.01	21.42	5.23	57.32
avNumAuth	9.01	8.60	0.08	9.13	8.30	0.22	8.72	8.06	6.76	9.81
sumPRI	1065.3	405.6	0.78***	875.2	370.8	0.57**	615.5	272.5	84.8	718.8
sumPRI*IF	6558.8	2154.1	0.94****	5378.0	1816.0	0.73***	3555.9	1468.0	641.6	3844.9
Mean			0.59			0.47				

B[b]

Parameter	Lea M	Lea F	Diff. 1 index	Sci M	Sci F	Diff. 2 index	Diff. 3 index	Diff. 4 index
Sample size	14	7	14 &7	18	23	18 & 23	14 & 18	7 & 23
ASI99	0.71	0.29	0.46	0.33	0.04	0.60*	0.44	0.83*

Table 2 continued

B[b]

Parameter	Lea M	Lea F	Diff. 1 index	Sci M	Sci F	Diff. 2 index	Diff. 3 index	Diff. 4 index
ASI95	2.14	1.43	0.24	1.50	0.48	0.54*	0.21	0.83*
ASI75	7.93	4.43	0.40	5.00	1.87	0.46	0.29	0.75*
ASI50	12.36	7.14	0.47	6.22	2.78	0.43	0.51	0.95**
NumP	21.36	11.57	0.58	8.06	4.83	0.36	0.82**	1.18**
sumC	510.4	238.4	0.63	230.1	79.6	0.52	0.61*	1.33***
avC	21.04	26.64	−0.41	19.14	14.33	0.33	0.13	0.86*
sumIF	145.3	61.5	0.81*	46.3	24.1	0.50	1.05***	1.70****
avIF	7.40	7.10	0.08	6.17	5.64	0.14	0.30	0.43
av%75	0.29	0.37	−0.35	0.42	0.29	0.33	−0.35	0.26
h-Index	10.3	6.7	0.56	5.22	3.04	0.44	0.71	1.52***
avPRI	59.24	58.00	0.10	58.03	49.61	0.36	0.06	0.36
avPRI*IF	407.7	357.4	0.32	313.5	277.3	0.18	0.57	0.37
avC/y	5.73	5.90	−0.06	5.22	3.76	0.35	0.12	0.67
sumC/y	125.2	54.3	0.70	57.5	20.1	0.55*	0.62*	1.22***
avNumAuth	8.99	8.77	0.06	9.24	8.16	0.27	−0.07	0.16
sumPRI	1255.4	685.1	0.53	579.5	275.2	0.42	0.60	1.03***
sumPRI*IF	8183.1	3311.9	0.80*	3195.7	1361.3	0.50	0.81**	1.17**
Mean			0.33			0.40	0.41	0.87

[a] *Lea* leader, *Sci* scientist, *F* woman, *M* man, *Lea F* woman leader, *Lea M* man leader, *Sci F* woman scientist, *Sci M* man scientist, *Me* median, *Q1* and *Q3* are quartiles. DI $(x_1 - x_2)$/SD. Significance is indicated by *asterisks*: $*p < 0.10$, $**\ p < 0.05$, $***p < 0.01$, $****p < 0.001$. For the meaning of line title see Table 1

[b] *Lea* leader, *Sci* scientist, *F* woman, *M* man, *Lea F* woman leader, *Lea M* man leader, *Sci F* woman scientist, *Sci M* man scientist. *Diff. 1* Lea M–Lea F, *Diff. 2* Sci M–Sci F, *Diff. 3* Lea M–Sci M, *Diff. 4* Lea F–Sci F. Other designations are the same as in Table 2A

The most expressed differences between "leaders" and "scientists" are seen in the following characteristics: sumIF (DI $= 1.18$, $p < 0.001$), NumP (DI $= 0.98$, $p < 0.001$), h-index (DI $= 0.96$, $p < 0.001$), sumPRI*IF (DI $= 0.94$, $p < 0.001$). The same parameters differentiate men and women as well, though the highest contrasts are in sumPRI*IF (DI $= 0.73$, $p < 0.01$) and sumC (DI $= 0.67$, $p < 0.01$).

Table 3 gives the correlation matrix for all the parameters considered. One can see that the parameters directly related to the number of papers (sumC, h-index, sumIF, etc.) are very highly correlated. The parameters reflecting impact (avC, avIF, avPRI) are weakly correlated or not correlated at all, though for many of them the correlation coefficients are significant, being above the critical value of 0.242 ($p < 0.05$, df $= 64$).

An important conclusion from these data is that the main parameter, characterizing the productivity of researchers is the NumP. All the other productivity indexes are highly correlated with the NumP and are determined by it. The most direct measure of impact seems to be the sumC (sum of citations over all the published papers), which is very highly ($r > 0.90$) correlated with other summary statistics: NumP, h-index, sumIF, sumPRI, and sumC/y. The parameters characterizing the quality of papers (rather than the number of them) poorly differentiate the compared groups and are weakly correlated with the NumP. These are avC, avIF, avPRI, and avC/y. This probably means that papers of the majority of the authors considered, both the leaders and regular scientists, men and women are of similar scientific level and exert similar impact on colleagues.

Figure 1 shows average "raw" citation numbers and those corrected for the age of the paper. One can see there is a considerable, clearly expressed time trend in "raw" citedness: the older the paper, the higher its citation frequency. Unlike "raw" citedness the age-corrected citedness is more or less similar in all publication years. Nevertheless, the "raw" citation values and those corrected for age strongly correlate (see Tables 3, 5) and the sizes of contrasts between the groups when using "raw" or corrected (sumC and sumC/y, avC and avC/y) values are not very different (see Tables 2A, B). This fact may be explained by the uniform distribution of paper publication times among the authors of all the compared groups.

As we mentioned above, our 2nd goal was to see how the PRI and ASI, which two of us (Pudovkin and Garfield 2009) suggested in 2009, perform on this data set. These indexes were introduced with the goal of finding alternative means for locating potentially impact papers and promising authors, not directly related to their "raw" citedness. It is well-known that in different science fields citation patterns and citation intensity differ, sometimes very significantly. PRI shows the citation status of a paper among papers by peers (published in the same source in the same year). Thus, PRI is a direct measure of impact judged by peers and does not require any citation thresholds or benchmarks (see the Appendix for explanation of these indexes).

Table 4 shows 11 papers with PRI $\geq$99. These papers are published in nine different journals with IFs ranging from 0.498 to 5.767. It is evident, that if one just looks at the 11 most-cited papers, nine would be missed. For instance, the 2008 paper in the Journal of periodontology is cited only 12 times. It is 142nd when sorted by "times cited" (see column "Rank by cites"), but its PRI $= 100$, which means it is the most highly cited paper among the papers of this journal in 2008. The 2007 paper in the Zeitschrift fur rheumatologie is cited only six times (184.5th when sorted by "times cited"), and again its PRI $= 100$. Thus, PRI allows one to locate papers of high citation status within a set of related papers (within a journal of a certain year or in a volume of a conference proceedings), even though its "raw" citation number is rather low.

Table 3 Product-moment correlation coefficients among 18 parameters characterizing productivity of 66 scientists

		f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v
b	*h*-Index	62	83	93	97	96	96	37	90	02	29	−06	27	23	30	96	97	97
f	ASI99	×	78	63	62	62	64	36	51	−07	29	−04	26	16	43	68	64	59
g	ASI95		×	92	89	81	85	35	62	−18	40	−09	36	11	34	86	88	77
h	ASI75			×	98	90	93	32	75	−14	40	−11	35	14	26	92	97	88
i	ASI50				×	96	95	30	84	−10	34	−13	32	15	23	95	99	94
j	NumP					×	92	23	92	−07	23	−12	21	10	16	92	98	94
k	sumC						×	46	90	03	32	−06	31	29	38	1.00	95	98
l	avC							×	34	38	49	03	55	80	93	47	29	41
m	sumIF								×	20	14	03	15	29	27	90	87	96
n	avJIF									×	−24	59	−15	67	37	04	−10	09
o	av%75										×	−09	76	29	51	33	31	27
p	NumAuth											×	−10	29	29	−04	−12	−03
q	avPRI												×	53	61	32	30	27
r	avPRI*IF													×	82	30	14	31
s	avC/y														×	41	23	34
t	sumC/y															×	95	98
u	sumPRI																×	95
v	sumPRI*IF																	×

For economy of space zeros and decimal points are omitted. Thus, the entry for the correlation coefficient between *h*-index and ASI99 should read 0.62. Critical values for the coefficients are 0.242 ($p < 0.05$) and 0.314 ($p < 0.01$), df = 64

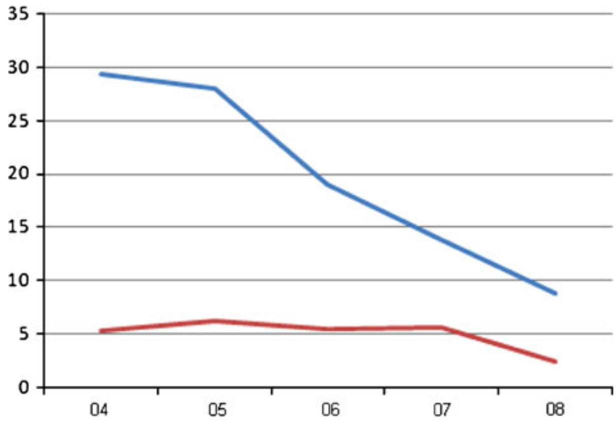


Fig. 1 Average citation rate of papers in 2004–2008. *Upper line* is "raw" average citation rate per paper ($avC_y = sumC_y/NumP_y$), *lower line* is the average citation rate corrected for years after publication ($avC/y_y = avC_y/(PY2 + 1 - PY1 - 0.5)$, where $sumC_y$ is the sum of citation received by 2010 by papers of the year "y", $NumP_y$ is the number of papers in the year "y", PY2 is the full year when citation numbers were retrieved, PY1 is the publication year of the paper. Subtraction of 0.5 allows to bring the data to the middle of the year. The citation data were retrieved from WOS in the beginning of 2010, thus the "full year" is 2009. For example, the value for papers of 2004 will be $(2009–2004 + 1 - 0.5) = 5.5$. In the ordinate is the average citation rate, in the abscissa is the publication year of papers

Table 4 Eleven papers with PRI ≥99

Journal title	Year	Cites	Rank by cites	PRI	IF	Num of auth	PRI*IF	Cites per year
Arthritis Research and Therapy	2006	70	18	100	3.801	6	380.1	20.0
British Journal of Dermatology	2007	51	32	100	3.503	11	350.3	20.4
Molecular and Cellular Endocrinology	2007	43	43	100	2.971	4	297.1	17.2
European Journal of Immunology	2008	42	46.5	100	4.865	10	486.5	28.0
Current Opinion in Rheumatology	2008	32	63.5	100	4.689	2	468.9	21.3
Journal of Periodontology	2008	12	142	100	1.961	12	196.1	8.0
Zeitschrift fur Rheumatologie	2007	6	184.5	100	0.651	5	65.1	2.4
Annals of the Rheumatic Diseases	2004	101	8	99	3.916	5	387.7	18.4
Annals of the Rheumatic Diseases	2006	89	10	99	5.767	6	570.9	25.4
European Journal of Immunology	2007	59	25	99	4.662	14	461.5	23.6
Hautarzt	2006	9	159	99	0.498	2	49.3	2.6

Table 5 gives the correlation values between some characteristics of the papers. It can be seen that the highest correlation of PRI is observed with cites/y (cites per year, 0.57, $p < 0.001$), and with "raw" cites (0.50, $p < 001$). One of the presumptive drawbacks of the PRI, on which reviewers commented, was a possible negative correlation of PRI with the journal IF: the lesser IF, the higher probability that a moderately cited paper will have a high PRI. Actually, the negative correlation is observed, but it is very small and non-significant (-0.06, $p > 0.05$). To further investigate this problem of biased values of PRI

Table 5 Product-moment correlation coefficients among six parameters characterizing 313 papers published by the staff of the DRFZ in 2004–2008

	Parameter	a	b	c	d	e	f
a	Cites	×	50	45	26	71	89
b	PRI		×	–06	11	43	57
c	IF			×	29	76	52
d	NumAuth				×	30	31
e	PRI*IF					×	82
f	Cites/y						×

Cites is the number of citations a paper has received by 2010. Cites/y is the number of cites per year: Cites/y = Cites/(PY2 + 1 − PY1 − 0.5), where PY2 is the full year when citation numbers were retrieved (2009 in our case), PY1 is the publication year of the paper

Zeros and decimal points are omitted. Critical values are 0.113 ($p < 0.05$) and 0.148 ($p < 0.01$); df = 311

(related to small IFs), we compared IF values of 35 papers with PRI ≥95 and the 278 papers with PRI <95. Student t-test turned out to be non-significant ($p = 0.182$). Thus, PRI appears to be a reliable measure of citation status, not significantly biased by smaller IF.

Interestingly, a high correlation (0.82, $p < 0.001$) is observed between PRI*IF and cites/y. This observation provides some basis for using this parameter (PRI*IF) in the evaluation of scientists' performance. Another point, though seemingly unimportant but of some interest, is the correlation of PRI with the number of authors. It is not strong but quite significant (0.11, $p < 0.01$). More strongly the number of authors correlates with IF (0.29, $p < 0.01$). The latter observation might be partly explained by the higher ambitions of multi-authored collectives—they are willing to take the chance of sending their papers to higher impact journals.

ASI99 and ASI95 allow identification of authors who do not have top or even high productivity indexes, but who did publish papers which were highly appreciated by their peers, that is by the readers of the journal in which the author has published the particular paper.

Table 6 illustrates this point. It shows citation data for 13 authors with the highest ASI99, ASI95. One can see, that among the 13 authors whose ASI99 ≥1 (that is who have at least one paper with PRI ≥99) there are some with rather small overall productivity indexes. For instance, the author FS has published only four papers from 2005–2008, but two of them have PRI ≥99. His other indexes are not high—h-index 4, which ranks at 30.5 from the top, the rank by sumIF = 35. The same refers to DJ, who published only four papers, but among them there is one with PRI ≥99. There is another person, WM, whose ASI99 ≥1, who is significantly lower ranked by other indexes. ASI95 reveals two more persons, HD and MK, who published two papers each of a high citation status. Their ASI95 = 2. This index shows also FS, who was revealed by ASI99; his ASI95 = 4.

Thus, the ASI99 and ASI95 allows one to find the authors, whose overall performance is not high, but who nevertheless published some papers attracting the attention and interest of their peers. Possibly, these persons are young and have only begun their scientific career, but nevertheless have published good papers, demonstrating their research potential.

Considering the publications of the DRFZ as a whole we can say that they are above the global average level: their average PRI is 58.5, while the global average would be about

Table 6 Authors with high values of ASI99, ASI95, and ASI75

Author	ASI99	Rank *h*-index	Rank NumP	Rank sumC	Rank sumIF	Rank sumC/y	Gender	Position	NumP	sumC	*h*-Index
S J	3	1	1	1	2	1	m	lea	64	1,709	26
B F	2	8	5	11	7	11	m	lea	40	526	13
D T	2	5.5	6.5	8	9	7	m	lea	28	623	15
F S	2	30.5	37	15	35	14	m	sci	4	241	4
H A	2	5.5	9	5	8	5	m	lea	24	913	15
H J	2	12.5	15	6	12	6	m	sci	14	757	10
D J	1	39.5	37	33	38	32	w	sci	4	73	3
L J	1	2	4	2	3	2	m	sci	41	1,469	25
R G	1	14.5	11	26	18	24	w	lea	20	149	9
R M	1	3	3	4	6	4	m	sci	42	1,050	21
S I	1	17.5	19	29	31	23	mow	sci	10	125	6
T A	1	10.5	8	12	10	12	m	lea	27	513	12
W M	1	30.5	29	23	25	22	w	lea	6	167	4
	ASI95										
S J	12	1	1	1	2	1	m	lea	64	1,709	26
L J	9	2	4	2	3	2	m	sci	41	1,469	25
R M	6	3	3	4	6	4	m	sci	42	1,050	21
H A	5	5.5	9	5	8	5	m	lea	24	913	15
H J	5	12.5	15	6	12	6	m	sci	14	757	10
Z A	5	8	6.5	10	11	10	w	lea	28	608	13
B F	4	8	5	11	7	11	m	lea	40	526	13
F S	4	30.5	37	15	35	14	m	sci	4	241	4
T A	4	10.5	8	12	10	12	m	lea	27	513	12
D T	3	5.5	6.5	8	9	7	m	lea	28	623	15
H H	3	12.5	13	13	16	13	w	sci	16	356	10
H D	2	17.5	17	21	22	20	w	sci	12	203	6
M K	2	30.5	15	34	39	38	w	sci	14	66	4

"*lea*" means "leader", "*sci*" means "regular scientist", "*m*" means "man", "*w*" means "woman", "*mow*" means the gender is not asserted. Other column heads are explained in Table 1

50.2 (for details see the Appendix), the difference is significant, $p < 0.001$; 183 papers of the 313 have PRI >50, which is more than expected $313/2 = 156.5$ (the difference is significant, $p < 0.025$). It is worthwhile to stress that the use of PRI allows one to compare the concrete set of papers with the global one, not employing any additional external benchmarks or thresholds as would be required when using other indicators (average citation rate, average IF, etc.). This is because the PRI inherently implies comparison with peer papers, those published in the same journal, in the same year (see Appendix for details).

The higher average citation status of "group leaders" vs. "regular scientists", both among men and women (see Tables 2A, B) speaks of a reasonable administering of staff resources: more productive (and presumably more experienced) persons are placed at leader positions.

It should be noted that the sum of IFs of the journals where the authors' papers are published is one of the most informative characteristics showing the difference both between "leaders" and "scientists" and between "male" and "female" subgroups. This gives support to the widespread, though often criticized practice of using this index in evaluation of candidates for grants, etc. This index may be preferable to other citation-based indexes when considering recent papers, which have not yet accumulated enough citations.

As was mentioned earlier, *h*-index, sumIF, sumC and NumP are strongly correlated (see Table 3). Though there are some discrepancies. For instance, among the authors we considered there are two persons, HA and DT, with the same *h*-index of 15. HA published fewer papers than DT (24 vs. 28), but overall citations of his papers is considerably higher (913 vs. 623). So, the question is which index is preferable? We believe that the indexes (*h*-index and sumC) are complementary. *h*-index characterizes stable publication output of moderately cited papers, while cumulative citation number (sumC) may reflect occurrence of a few serendipitous, highly cited papers.

It is also noteworthy that members of the compared subgroups have some co-authored papers—the overall number of papers we considered is 313, but the sum of the paper numbers over the authors is 660. One should also keep in mind that the 313 papers by the 66 authors we considered were co-authored by 1,045 persons, of which $1{,}045–66 = 979$ persons were from other domestic and foreign institutions.

Considering the compared groups and subgroups (Tables 2A, B) one should be aware that these groups are not homogenous: within each groups there are some very productive authors and also poorly productive ones. Thus, the group comparisons reflect only average tendencies. There is a very strong overlap in the index values between the compared groups. For instance, the strongest contrast is observed in sumPRI*IF between the subgroups: "female leaders" and "female scientists", the difference in this index being 1.70 ($p < 0.001$). But even in this case the overlap is considerable (Fig. 2). One should also keep in mind that our samples are small, indeed some are very small: the subgroup "female leaders" consisted of only seven persons.

Conclusions

These conclusions are valid only in relation to the dataset studied—66 researchers, of the DRFZ, group leaders and ordinary scientists, men and women, who have published 313 papers from 2004–2008.

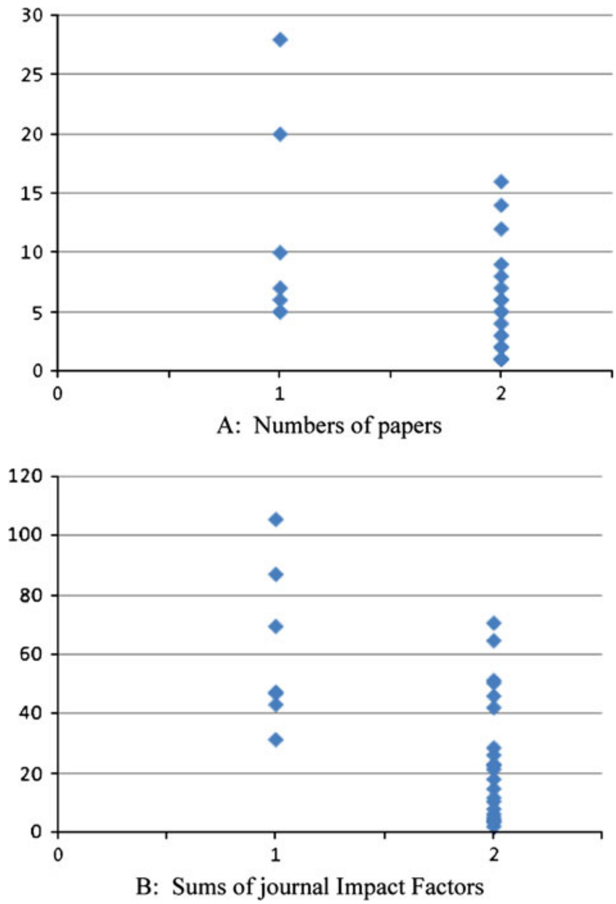


Fig. 2 Distribution of numbers of papers and sums of journal IF among women, leaders (1) and scientists (2)

1. Leaders are more prolific and more cited than the regular scientists.
2. Men are more productive than women.
3. The greatest contrast is observed between female leaders and female regular scientists.
4. The parameters in which the differences between groups are more pronounced are the number of papers, cumulative number of cites to these papers, *h*-index, sum of IF of the journals in which the papers are published.
5. The latter fact supports the use of the sum of the IF of the recent publications in grant applications and evaluations for promotions. This characteristic may be preferable in evaluating recent publications, for which citation data is not yet available.
6. ASI99 and ASI95 can be instrumental in revealing promising young authors.

Acknowledgments Part of this work by one of the authors (Kretschmer, H.) was supported by the 7th framework program by the European Commission, SIS-2010-1.3.3.1. Project full title: "Academic Careers Understood through Measurement and Norms", Project acronym: ACUMEN.

Appendix: Calculation of the PRI and the ASI

The procedure for obtaining the PRI and ASI is described in Pudovkin and Garfield (2009). It is a two step process. It requires that we first obtain the PRI for each of the individual papers an author has published, and then calculate the ASI, which is based on PRI values for all the author's papers. The PRI for each paper is based on the citation rank of the paper among the papers published in the same journal in the same year. In other words, the comparison is made among the related papers of the target one published within the same specialty journal or any topical group of papers one may aggregate by various methods, provided the papers are of the same age as the other papers under comparison. Thus, PRI also may be applied to papers published in multi-authored books, proceedings volumes, or other topical collections of papers.

To retrieve the necessary data an author search is conducted within WOS to find all the papers of the specified author covered by WOS. For each paper one makes a journal search for a specific year and retrieves all the papers published by that journal in the same year. Then one clicks on the "citation report" button in WOS. This option sorts the papers by citation frequency and calculates the average citation rate. To calculate the PRI one needs the citation rank of the paper and the number of papers in the year set of the journal. Both are provided by the "citation report" option.

$$\mathrm{PRI} = (N - R + 1)/N * 100 \tag{1}$$

where N is the number of papers in the year set of the journal, R is the descending citation rank of the paper (among the papers of the journal published in the year of the target paper). In case of ties (several papers having the same citation frequency), each of the tied values is assigned the average of the ranks for the tied set. Thus, if a target paper is the most cited paper in a journal in a year, its PRI = 100.

Consider the paper in the 4th line of Table 7. It contains information on a paper published in 2006 in the journal Arthritis Research and Therapy. By March of 2010 it has been cited 70 times. The overall number of papers published in this journal in 2006 is 228. This paper is the most cited among the 228 papers of this journal (published in 2006). Thus, its citation rank is 1. Entering these values into the formula (1) we obtain PRI = (228 − 1 + 1)/228*100 = 100. Another example: a paper in the journal Arthritis and Rheumatism, 2005 (the 2nd line). It has been cited (by March 2010) 133 times, being the top 14th by "raw" citation number. Entering the data into the formula we obtain PRI = (437 – 14 + 1)/437*100 = 97.02, which is rounded up to 97.

The average PRI value over all the papers of a journal (within a year set) would be about 50.5. It would depend on the number of papers in the journal: for larger journals with the number of papers >100 it will be closer to 50, while for smaller journals the average will be higher. Exact value of average PRI = (50 + $1/N$*100/2), where N is the number of

Table 7 Data for calculation of PRI

Journal title	Publication year	Number of cites	Number of papers	Citation rank	PRI
Journal of Experimental Medicine	2004	229	320	12	97
Arthritis and Rheumatism	2005	133	437	14	97
Nature Reviews Immunology	2006	114	82	24	72
Arthritis Research and Therapy	2006	70	228	1	100

papers in the year set of the journal. For instance, for a journal with 50 papers the average PRI will be 51.0, for a journal with 20 papers it will be 52.5. For our data set the expected average PRI is 50.2 as the median number of papers in the journals considered is 275. This property of the average PRI allows to directly see (without any external benchmarks or thresholds) whether an author (or a group of authors) performs better than his peers on average.

ASI is the number of papers of an author, which have PRI_X at (or higher) the percentile threshold of X. Thus, the ASI99 is the number of papers, which PRI is equal or higher than 99. Similarly, ASI95 is the number of papers with PRI $\geq$95.

References

Bar-Ilan, J. (2008). Informetrics at the beginning of the 21st century—a review. *Journal of Informetrics, 2*(1), 1–52.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, 102*(46), 16569–16572.

Ioannidis, J. P., Patsopoulos, N. A., Kavvoura, F. K. et al. (2007). International ranking systems for universities and institutions: a critical appraisal. *BMC Medicine*, *5*(30).

Pudovkin, A. I., & Garfield, E. (2009). Percentile rank and author superiority indexes for evaluating individual journal articles and the author's overall citation performance. *Collnet Journal of Scientometrics and Information Management, 3*(2), 3–10.

Sanz-Casado, E., Iribarren-Maestro, I., Garcia-Zorita, C. et al. (2009). Are productivity, impact and visibility indicators appropriate for measuring the quality of research conducted in universities? In B. Larsen and J. Leta (Eds.), *Proceedings of ISSI 2009—12th International Conference of the International Society for Scientometrics and Informetrics* (Vol. 1. pp. 286–290).

Tijssen, R. J. W., van Leeuwen, T. N., & van Wijk, E. (2009). Benchmarking university-industry research cooperation worldwide: Performance measurements and indicators based on co-authorship data for the world's largest universities. *Research Evaluation, 18*(1), 13–24.

Wallin, J. A. (2005). Bibliometric methods: Pitfalls and possibilities. *Basic and Clinical Pharmacology and Toxicology, 97*(5), 261–275.