

The top-cited research works in the Science Citation Index Expanded

Yuh-Shan Ho

Received: 10 April 2012 / Published online: 1 September 2012
© Akadémiai Kiadó, Budapest, Hungary 2012

Abstract This study aimed to identify and analyze the characteristics of the top-cited articles published in the Science Citation Index Expanded from 1991 to 2010. Articles that have been cited more than 1,000 times since publication to 2010 were assessed regarding their distribution in indexed journals and categories of the Web of Science. Five bibliometric indicators were used to evaluate source institutions and countries. A new indicator, the Y-index, is proposed to assess publication quantity and the character of contribution to articles. We identify 3,652 top-cited articles with 71 % originating from US. The fourteen most productive institutions were all located in US. *Science*, *Nature*, *New England Journal of Medicine*, and *Cell* hosted the most cited publications. In addition, the Y-index was successfully applied to evaluate the publication character of authors, institutions, and countries.

Keywords Web of Science · SCI-Expanded · Bibliometric · Most cited publication · Y-index

Introduction

Garfield (1955) in his classic paper “citation indexes for science: a new dimension in documentation through association of ideas” may be said to have launched contemporary scientometrics. Since 1964 the Science Citation Index (SCI) has been a leading tool in indexing (Garfield 1964). The SCI is a valuable sociometric tool for historians and sociologists (Garfield 1970) and a unique and necessary tool for scientific work (Malin 1968). The SCI presents data in the form of what has been published and the associated citations of publications. It has become one of the most widely and frequently used databases for searching literature and evaluating research performances. Basically, the SCI offers two main sources of information: “what has been published”? and “what are the relationships among these publications”? The citation index of an article is not a direct measure of its

Y.-S. Ho (✉)
Trend Research Centre, Asia University, Taichung 41354, Taiwan
e-mail: ysho@asia.edu.tw

quality or importance; it is a measure of recognition that may suggest its visibility or impact on the scientific community (Furlan and Fehlings 2006). In earlier years, the SCI has been applied to the analysis of the most cited life sciences (Garfield 1989), physical sciences (Brush 1990) publications, and the most cited Soviet papers (Garfield 1990).

Thus, for example in the last decade, the most cited articles in medical research fields were orthopaedic surgery (Lefavre et al. 2011), ophthalmology (Ohba and Nakao 2010), critical care medicine (Rosenberg et al. 2010), urology (Hennessey et al. 2009), pediatric surgical research (Celayir et al. 2008), occupational medicine (Gehanno et al. 2007), periodontology (Nieri et al. 2007), traumatic spinal cord injury (Furlan and Fehlings 2006), and trauma (Ollerton and Sugrue 2005).

General conclusions publishing behaviour include the observation that the top-cited papers were published in high impact factor journals. North America was most active in producing top-cited papers. Researchers for instance, Celayir et al. (2008), Rosenberg et al. (2010), Ohba and Nakao (2010) investigated the research focus of the top-cited papers. Furthermore, it was found that Nobelists are consistently highly cited while only a small percentage of most-cited authors win the prize. It would be expected that a large percentage of the latter are elected to national academies of science (Garfield and Well-jamsdorof 1992).

This corresponds with the observation that equal credit is not given to all contributors of a publication. At the individual level, a non-alphabetical name order sends a clear signal to the market that the author who is listed first has actually contributed more (Engers et al. 1999). The first author is the person who contributed most to the work and writing of the article (Gaeta 1999). It has also been mentioned in guidelines on authorship of medical papers that the first author should have made major contributions in conception of the work represented by the article, design of the work, analysis and interpretation of data or other evidence presented in the article as well as drafting the article or revising it for critically important content (Huth 1986). The corresponding author is perceived as the author contributing significantly to the article independently of the author position (Mattsson et al. 2011). The corresponding author supervised the planning and execution of the study and the writing of the paper (Burman 1982). At the country or institutional level, the country or institution of the corresponding author might be a home base of a study, or origin of the paper.

In this study, all journal articles with more than 1,000 total citations since publication to 2010 were selected as top-cited research works and analyzed with regard to citation histories, total citation, and citation in 2010, journals, and Web of Science categories. Top-cited publications with authors address information were further analyzed with a new indicator the Y-index that was developed and used to evaluate contributions of individual authors, institutions, and countries.

Methodology

The information on documents used in this study is based on the Science Citation Index Expanded (SCI-Expanded) database of the Thomson Reuters Web of Science. According to Journal Citation Reports (JCR) of 2010, SCI-Expanded indexes 8,073 journals with citation references across 174 scientific disciplines in science edition. The journal index of the Web of Science was last updated on 29 February 2012. All the papers published in the last 10 years of the twentieth century and the beginning 10 years of twenty first century, were collected. The citation frequencies for each of the top-cited papers per year were

collected from papers published to 2010. The total citations of a paper in recent year (2010) was recorded as *C2010* and the total number of citations of an article from its publication to 2010 was recorded as *TC2010* (Wang et al. 2011; Chuang et al. 2011). The articles with $TC2010 \geq 1,000$ were selected as top-cited articles. The records were downloaded into spreadsheet software, and additional coding was manually performed to obtain the frequency distributions and percentages. Articles originating from England, Scotland, Northern Ireland, and Wales were reclassified as being from the United Kingdom (UK). USSR and Russia were reclassified as being from Russia. Czechoslovakia and Czech Republic were also reclassified as being from Czech Republic. Yugoslavia and Croatia were also reclassified as being from Croatia. Similarly, Articles from Hong Kong published before 1997 were included in the China category. Collaboration type was determined by the addresses of the authors, whereas, the term “country independent article” was assigned if the researchers’ addresses were from the same country. The term “internationally collaborative article” was designated to those articles that were coauthored by researchers from multiple countries (Chiu and Ho 2005). The term “institution independent article” was assigned if the researchers’ addresses were from the same institution. The term “inter-institutionally collaborative article” was assigned if authors were from different institutions (Li and Ho 2008). The impact factor of a journal was determined for each document as reported in the JCR 2010.

Y-index

It has been accepted that the most important positions are the first and the last, whom very often is the corresponding author (Zuckerman 1968; Costas and Bordons 2011). The first author contributed most to the work, including conducting research and writing of the manuscript (Herbertz and Müller-Hill 1995; Riesenbergs and Lundberg 1990). It was also noticed that the corresponding author supervised the planning and execution of the study and the writing of the paper (Burman 1982). It has been found that an increase number of authors in a paper is more likely to precipitate various unethical authorship practice including gift authorship (Slone 1996; Dotson and Slaughter 2011). In addition, gift or honorary authorship is defined as the inclusion as author of an individual who has not contributed adequately to the project (Bennett and Taylor 2003; Singh 2009). However, the honorary authorship is still regarded as a minor digression and the honorary or gift authorship is unacceptable in the *Lancet* (2008). In this study, the Y-index is related to important positions which are the numbers of first author publications (FP) and corresponding author publications (RP). In general, only one parameter was included in indexes, for example h-index (Hirsch 2005), g-index (Egghe 2006), A-index (Jin 2006), R-index (Jin et al. 2007), and AR-index (Jin et al. 2007). The construction of the Y-index with two parameters (j, θ), is an attempt to assess both the publication quantity and character of contribution as a single index. The Y-index is defined as:

$$j = \sqrt{FP^2 + RP^2} \tag{1}$$

$$\theta = \tan^{-1} \left(\frac{RP}{FP} \right) \tag{2}$$

j indicates publication quantity with important author positions (first and corresponding authors) articles only. It was calculated by using numbers of first authored articles and

corresponding authored articles as the Eq. (1). When one had larger j it could be also found one's Y-index located far away from original of the polar coordinates. It means that one published more articles as "important author". In order to make sure where Y-index would be in the polar coordinates, another parameter θ is necessary. θ is a publication character constant, that differentiates its nature of leadership role. It introduces distribution of the numbers of the first authored articles and the corresponding authored articles. When the numbers of the first authored articles and the corresponding authored articles are the same, Y-index is located in the 45 degree (0.7854 rad) line. Thus θ could be calculated by using Eq. (2). Then when $\theta > 0.7854$, means one published more corresponding author papers and $\theta < 0.7854$, means one published more first author papers. When $\theta = 0$, j number of first author papers and $\theta = \infty$, j number of corresponding author papers.

In the SCI-Expanded database, the corresponding author is labeled as the reprint author. In our study this person is referred to as the corresponding author. In a single author article where authorship is not specified, the author is classified as the first author and the corresponding author. The Y-index was calculated and was applied to evaluate country, institution, and individual author publication characters. In total 3,022 documents with both first and corresponding authors were analyzed using the Y-index.

Results and discussion

In total 21,066,849 papers were published in SCI-Expanded from 1991 to 2010. There were 3,652 papers with $TC2010 \geq 1,000$. These papers consist of 9 document types—articles (2,541) dominate with a comprising 70 %, followed distantly by reviews (951; 26 %). English is the only one language used for the top-cited papers. Finally, the journal articles (2,541) were extracted from the 3,652 documents for subsequent analyses.

Table 1 shows the number of total citations since papers publication to 2010 ($TC2010$) were further considered for the citations per publication (CPP) which was similarly for most document types but software reviews had highest CPP (4,128).

Table 1 Document type distribution

Document type	TP (%)	TC2010 (%)	CPP
Article	2,541 (70)	4,441,218 (69)	1,748
Review	951 (26)	1,666,201 (26)	1,752
Proceedings paper article	50 (1.4)	80,764 (1.3)	1,615
Note	49 (1.3)	92,347 (1.4)	1,885
Editorial material	36 (1.0)	61,656 (1.0)	1,713
Letter	15 (0.41)	27,070 (0.42)	1,805
Software review	8 (0.22)	33,026 (0.52)	4,128
Addition correction	1 (0.027)	1,618 (0.025)	1,618
Database review	1 (0.027)	1,346 (0.021)	1,346
Total	3,652 (100)	6,405,246 (100)	1,754

Effect of time on citation analysis

Figure 1 shows geographical distribution of these top-cited articles. North America, West Europe, and Japan were the main production area. Seventy percent of articles identified were published from 1991 to 1998 (Fig. 2). It also shows that the years from 2007 to 2009, although there were fewer top-cited articles, on average they have higher citations (CPP > 2,000) than earlier years. In 2008, there were only three top-cited articles. Article “a short history of SHELX” published in *Acta Crystallographica Section A* by Sheldrick (2008) had $TC_{2010} = 15,241$. This article had the highest yearly citation in 2010 with 6,826 citations. “CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice” (Thompson et al. 1994) with $TC_{2010} = 31,799$, “density-functional thermochemistry. III. The role of exact exchange” (Becke 1993) with $TC_{2010} = 29,822$, and “gapped BLAST and PSI-BLAST: a new generation of protein database search programs” with $TC_{2010} = 26,578$ (Altschul et al. 1997) had highest TC_{2010} in SCI-Expanded from 1991 to 2010. “Structure validation in chemical crystallography” published in *Acta Crystallographica Section D-Biological Crystallography* (Spek 2009) and “Cancer Statistics 2009” published in *CA-A Cancer Journal for Clinicians* (Jemal et al. 2009), were the latest top-cited articles with TC_{2010} 2,613 and 1,945 respectively.

Table 2 presents data on the 12 articles cited more than 10,000 times. Out of these 12 articles, three were internationally collaborative publications and nine were independent publications. The USA published five independent articles followed by Germany with two articles and one for Canada and Japan respectively. The journals in which these articles published were *Nucleic Acids Research* (IF = 7.836) in the field of biochemistry and molecular biology with three articles, followed by *Acta Crystallographica Section A* (IF = 54.333) in the field of crystallography with two articles, and one article for each of *Methods* (IF = 4.527), *Acta Crystallographica Section D-Biological Crystallography* (IF = 3.326), *Nature* (IF = 36.104), *Bioinformatics* (IF = 4.877), *Physical Review Letters* (IF = 7.621), *Journal of Chemical Physics* (IF = 2.92), and *Medical Care* (IF = 3.183) respectively. Three authors had single author article including Becke from Canada, Sheldrick from Germany, and Iijima from Japan. Thompson from France who was the only authors published two first author articles.

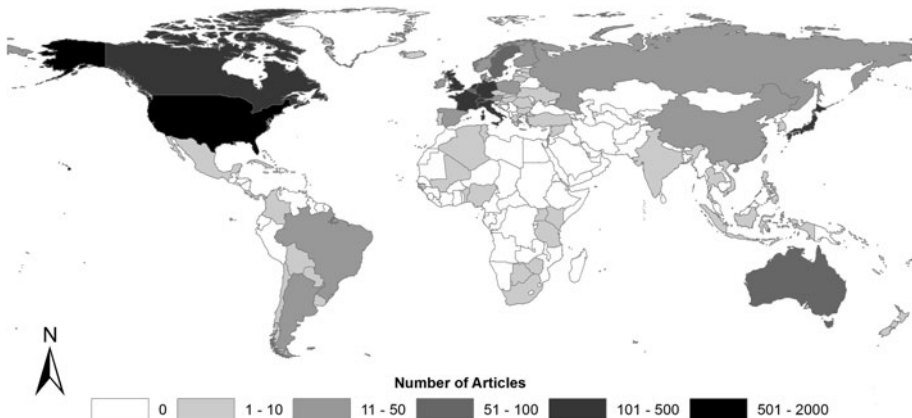


Fig. 1 Distribution of top-cited articles in the world

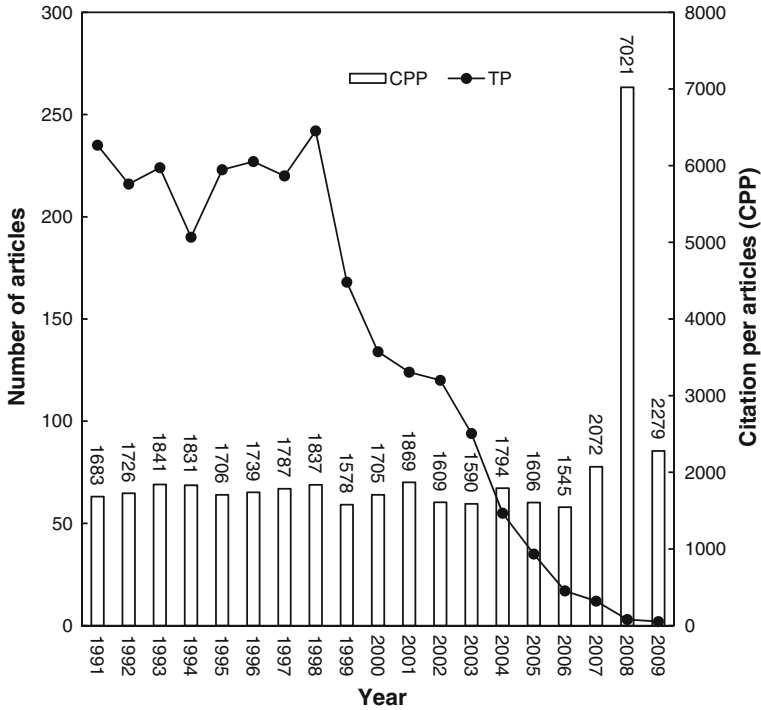


Fig. 2 Number of articles on the top-cited article and citation per articles

Table 2 Twelve most frequently cited articles in the SCI-Expanded (*TC2010* > 10,000)

<i>TC2010</i> (rank)	<i>C2010</i> (rank)	Article information
31,799 (1)	2,524 (6)	Thompson et al. (1994)
29,822 (2)	3,036 (4)	Becke (1993)
26,578 (3)	2,469 (7)	Altschul et al. (1997)
17,666 (4)	2,025 (8)	Thompson et al. (1997)
16,381 (5)	3,405 (2)	Perdew et al. (1996)
15,240 (6)	6,826 (1)	Sheldrick (2008)
12,978 (7)	817 (44)	Brünger et al. (1998)
12,035 (8)	397 (163)	Jones et al. (1991)
10,582 (9)	1,112 (28)	Iijima (1991)
10,497 (10)	2,573 (5)	Livak and Schmittgen (2001)
10,428 (11)	1,576 (16)	Posada and Crandall (1998)
10,304 (12)	1,069 (30)	Ware and Sherbourne (1992)

TC2010 number of citations till 2010, *C2010* number of citations in 2010

Journal and Web of Science category

The top-cited articles were published by 365 journals across 122 subject categories. Of these 365 journals, 187 (51 %) journals contained only one top-cited article; 68 (19 %)

journals contained two articles; and 31 (8.5 %) journals contained three articles. 2,491 articles were published in journals with impact factor information in JCR in 2010. Fifty-five percent of the top-cited articles were published in 10 journals with impact factors higher than 30. Half of the top-cited articles were published in five journals including *Science* (IF = 31.364) with 406 articles (16 %), followed by *Nature* (IF = 36.104) with 378 articles, *New England Journal of Medicine* (IF = 53.484) with 203 articles, *Cell* (IF = 32.401) with 197 articles, and *Proceedings of the National Academy of Sciences of the United States of America* (IF = 9.771) with 87 articles (Table 3).

As expected, top-cited articles were published in journals with the high impact factor, similar to the phenomenon in the study related to anesthetics (Baltussen and Kindler 2004). It was also noticed that leading journals attract the most-cited publications, which in turn

Table 3 Characteristics of top 20 journals with the top-cited articles

Journal	Rank	TP (%)	IF2010	Web of Science category
Science	15	406 (16)	31.377	Multidisciplinary sciences
Nature	9	378 (15)	36.104	Multidisciplinary sciences
New England Journal of Medicine	3	203 (8.0)	53.486	General and internal medicine
Cell	14	197 (7.8)	32.406	Biochemistry and molecular biology cell biology
Proceedings of the National Academy of Sciences of the United States of America	151	87 (3.4)	9.771	Multidisciplinary sciences
Lancet	11	77 (3.0)	33.633	General and internal medicine
JAMA-Journal of the American Medical Association	18	65 (2.6)	30.011	General and internal medicine
Physical Review Letters	238	53 (2.1)	7.622	Multidisciplinary physics
Journal of Chemical Physics	1,462	27 (1.1)	2.921	Atomic, molecular and chemical physics
Nucleic Acids Research	228	26 (1.0)	7.836	Biochemistry and molecular biology
Journal of Experimental Medicine	69	24 (0.94)	14.776	Immunology research and experimental medicine
Nature Genetics	8	23 (0.91)	36.377	Genetics and heredity
Nature Medicine	32	20 (0.79)	25.430	Biochemistry and molecular biology cell biology research and experimental medicine
Physical Review B	923	20 (0.79)	3.774	Condensed matter physics
Circulation	72	18 (0.71)	14.432	Cardiac and cardiovascular systems hematology peripheral vascular disease
Applied Physics Letters	884	18 (0.71)	3.841	Applied physics
CA-A Cancer Journal for Clinicians	1	17 (0.67)	94.333	Oncology
Journal of the American Chemical Society	182	17 (0.67)	9.023	Multidisciplinary chemistry
Journal of Biological Chemistry	427	17 (0.67)	5.328	Biochemistry and molecular biology
Journal of Molecular Biology	815	17 (0.67)	4.008	Biochemistry and molecular biology

Rank rank of impact factor in 8,073 journals in JCR in 2010, *TP* total number of top cited articles, *IF2010* impact factor in 2010

Table 4 Characteristics of the 29 countries with first author publications

Country	TP	TP R (%)	SP R (%)	CP R (%)	FPR (%)	RPR (%)	θ	$j R (j)$	$j R (j_s)$
USA	2,160	1 (71)	1 (70)	1 (76)	1 (63)	1 (64)	0.7911	1 (2715)	1 (2230)
UK	458	2 (15)	2 (8.3)	2 (35)	2 (9.4)	2 (9.4)	0.7872	2 (400.9)	2 (264.5)
Germany	282	3 (9.3)	3 (4.1)	3 (25)	3 (4.6)	3 (4.5)	0.7781	3 (193.8)	3 (131.5)
France	227	4 (7.5)	5 (2.4)	4 (23)	5 (3.5)	5 (3.5)	0.7760	5 (149.9)	5 (76.37)
Canada	212	5 (7.0)	6 (2.3)	5 (21)	6 (2.8)	6 (2.9)	0.7970	6 (121.6)	6 (74.95)
Japan	181	6 (6.0)	4 (4.0)	9 (12)	4 (3.8)	4 (3.7)	0.7765	4 (159.8)	4 (127.3)
The Netherlands	127	7 (4.2)	7 (1.3)	6 (13)	8 (1.6)	8 (1.6)	0.7749	8 (67.18)	7 (42.43)
Switzerland	124	8 (4.1)	7 (1.3)	7 (12)	7 (1.9)	7 (1.6)	0.7188	7 (74.41)	7 (42.43)
Italy	113	9 (3.7)	10 (0.93)	8 (12)	9 (1.6)	9 (1.3)	0.7051	9 (61.72)	10 (29.70)
Australia	96	10 (3.2)	9 (1.0)	11 (10)	10 (1.2)	11 (1.1)	0.7432	10 (50.25)	9 (32.53)
Sweden	93	11 (3.1)	12 (0.66)	10 (10)	11 (1.0)	10 (1.2)	0.8459	11 (46.75)	12 (21.21)
Belgium	69	12 (2.3)	13 (0.49)	13 (7.6)	13 (0.76)	13 (0.79)	0.8067	13 (33.24)	13 (15.56)
Spain	66	13 (2.2)	16 (0.27)	12 (7.8)	16 (0.50)	16 (0.40)	0.6747	16 (19.21)	16 (8.485)
Israel	58	14 (1.9)	11 (0.75)	15 (5.4)	12 (0.79)	12 (0.93)	0.8622	12 (36.88)	11 (24.04)
Denmark	56	15 (1.9)	15 (0.40)	14 (6.1)	14 (0.63)	15 (0.53)	0.6999	15 (24.84)	15 (12.73)
Austria	44	16 (1.5)	14 (0.44)	16 (4.4)	15 (0.60)	14 (0.69)	0.8622	14 (27.66)	14 (14.14)
Finland	39	17 (1.3)	16 (0.27)	17 (4.3)	17 (0.36)	18 (0.26)	0.6288	17 (13.60)	16 (8.485)
Norway	29	18 (1.0)	19 (0.13)	18 (3.4)	19 (0.23)	19 (0.23)	0.7854	19 (9.899)	19 (4.243)
China	27	19 (0.89)	21 (0.089)	19 (3.3)	18 (0.26)	17 (0.30)	0.8442	18 (12.04)	21 (2.828)
Russia	26	20 (0.86)	19 (0.13)	20 (3.0)	19 (0.23)	22 (0.13)	0.5191	20 (8.062)	19 (4.243)
India	18	22 (0.60)	18 (0.22)	23 (1.7)	21 (0.17)	20 (0.17)	0.7854	21 (7.071)	18 (7.071)
Poland	15	23 (0.50)	25 (0.044)	22 (1.8)	29 (0.033)	27 (0.033)	0.7854	29 (1.414)	25 (1.414)
Brazil	14	24 (0.46)	21 (0.089)	25 (1.6)	23 (0.10)	23 (0.10)	0.7854	23 (4.243)	21 (2.828)
Hungary	11	26 (0.36)	21 (0.089)	27 (1.2)	27 (0.066)	23 (0.10)	0.9828	26 (3.606)	21 (2.828)
South Korea	10	28 (0.33)	21 (0.089)	29 (1.0)	23 (0.10)	23 (0.10)	0.7854	23 (4.243)	21 (2.828)
New Zealand	9	29 (0.30)	25 (0.044)	29 (1.0)	23 (0.10)	23 (0.10)	0.7854	23 (4.243)	25 (1.414)

Table 4 continued

Country	TP	TP R (%)	SP R (%)	CP R (%)	FP R (%)	RP R (%)	θ	$j R (j)$	$j R (j_s)$
Singapore	8	31 (0.26)	25 (0.044)	33 (0.91)	29 (0.033)	27 (0.033)	0.7854	29 (1.414)	25 (1.414)
Taiwan	8	31 (0.26)	25 (0.044)	33 (0.91)	29 (0.033)	27 (0.033)	0.7854	29 (1.414)	25 (1.414)
Czech Republic	6	36 (0.20)	25 (0.044)	37 (0.65)	27 (0.066)	27 (0.033)	0.4636	28 (2.236)	25 (1.414)

TP total publications, *R* rank, *SP* single country publications, *CP* collaborative publications, *FP* first author publications, *RP* corresponding author publications, *j* Y-index parameter, publication intensity constant, θ Y-index parameter, publication character constant, j_s single country publication intensity constant

maintain the high impact factor of these journals (Schein et al. 2000). However, articles with $TC_{2010} > 1,000$ could be also found in journals with lower impact factors such as *European Transactions on Telecommunications* with $IF = 0.448$, *Applied Clay Science* with $IF = 2.303$, *Journal of the Electrochemical Society* with $IF = 2.42$ in category of telecommunications; *Optimization Methods and Software* with $IF = 0.794$ in category of software engineering computer science, operations research and management science, and applied mathematics; *AI Communications* with $IF = 0.837$ in category of artificial intelligence computer science; *Theoretical Computer Science* with $IF = 0.838$ in category of theory and methods computer science; and *Spatial Vision* with $IF = 0.883$ in category of biophysics.

Within the total 122 Web of Science categories, 65 categories (53 %) accounted for 0–5 top-cited articles, 14 categories (12 %) for 6–10 articles, 23 categories (19 %) for 11–30 articles, and 20 categories (16 %) for more than 30 articles. In particular, the four top categories including multidisciplinary sciences with 875 articles, biochemistry and molecular biology with 388 articles, general and internal medicine with 370 articles, and cell biology with 274 articles, took the majority of the total top-cited articles with a high percentage of 75 %. Thirteen categories had two top-cited articles, thirty-two categories had only one and fifty-two categories had none.

Publication performances: countries, institutions, and authors

In recent years, indicators of performance of first authors (Li and Ho 2008), institutions (Ho et al. 2010) and countries (Wang et al. 2010) were examined to compare research performances. Among 3,652 top-cited documents, 3,022 publications had both first authors and corresponding author information, including 1999 articles, 889 reviews, 49 proceedings paper articles, 33 notes, 32 editorial materials, 12 letters, 7 software reviews and one database review were analyzed for publications of countries, institutions, and authors.

The Y-index was used to analyze publication performance of the 3,022 top-cited publications including 2,256 (75 %) country independent publications from 29 countries and 766 (25 %) internationally collaborative publications from 62 countries. Table 4 shows the 29 countries which had independent publications, ranked according to the number of total top-cited papers published, in terms of number total papers, country independent papers, internationally collaborative papers, first author papers, and corresponding author's papers. Moreover the Y-index constants, θ , j , and their ranks, were also presented. United States (USA) was the most productive country in five indicators. It had strongest publication intensity with $j = 2,715$ while Poland, Taiwan, Singapore, and Czech Republic had the lowest j ($j = 1.141$). USA published more corresponding author papers than first author papers with $\theta = 0.7911$. Czech Republic, Russia, Finland, Spain, Denmark, Italy, Switzerland, Australia, The Netherlands, France, Japan, and Germany had $\theta < 0.7854$, that means these countries published more first author papers. In addition, Ireland ranked 21th in total top-cited publications, Mexico (25th), Chile (26th), Argentina (29th), Iceland (31st), South Africa (31st), Greece (35th), and Portugal (36th). No first author papers were published by these countries. The G7 countries (the US, UK, Canada, Germany, France, Italy, and Japan) had high productivity in top-cited publications, which included 2,814 (93 % of 3,022 top-cited publications). Domination in publication from mainstream countries was not surprising since this pattern has occurred in many medical related topics, for example patent ductus arteriosus (Hsieh et al. 2004), asthma in children (Chen et al. 2005), stem cells (Li et al. 2009a), *Helicobacter pylori*

(Suk et al. 2011), human papillomavirus (Lin et al. 2011), and desalination (Tanaka and Ho 2011).

Of the 3,022 publications with both first and corresponding author information in Web of Science, 1,347 (45 %) publications came from independent institutions, 1,675 (55 %) publications from inter-institutional collaboration. The inter-institutional collaboration rate of top-cited articles (55 %) was observed to be greater than with that of many fields with all related articles, for example 53 % of atmospheric simulation (Li et al. 2009b) and acupuncture research (Han and Ho 2011), 44 % of solid waste research (Fu et al. 2010), 37 % of desalination research (Tanaka and Ho 2011), 50 % of articles in water resources field (Wang et al. 2011) as well as some medical fields with respect to classic citation articles, for example 12 % of 100 top-cited articles in general surgical journals (Paladugu et al. 2002), 8 % of 100 ophthalmology class citations (Ohba et al. 2007); but smaller than fields for example 60 % of *Helicobacter pylori* research (Suk et al. 2011) and 62 % of global climate change (Li et al. 2011).

Results from several analytical methods for the analysis, including correlations, nonparametric tests (e.g., the Mann–Whitney test), and multidimensional scaling (MDS) pointed out that a majority of US universities were dominant in the center (Lee and Park 2012). Table 5 shows the top 20 institutions published at least 50 top-cited papers, ranked according to the total number of top-cited articles. Among the top 20 institutes, 18 (90 %) located in USA. Harvard University ranked first with 231 papers, followed by Stanford University (120 papers), and University of Texas (108 papers), and Massachusetts Institute of Technology (106 papers). Brigham & Womens' Hospital, Massachusetts General Hospital, and National Cancer Institute were three non-university institutions. These two hospitals had less independent publications. The two non-US institutions were University of Oxford and University Cambridge, ranked in 15th and 18th, respectively. Furthermore Duke University ranked 12th and Rockefeller University ranked 13th in independent publications. Duke University, American Cancer Society, and Salk Institute for Biological Studies, and Osaka University, all ranked 17th in first author publications. Salk Institute for Biological Studies also ranked 15th in corresponding author publications.

Harvard University not only ranked top in total articles but also the top one in j . It has been reported that Harvard University was ranked as one, which was re-coded as 100 for easy interpretation (Lee and Park 2012). However, the rank of total top-cited publication and the rank of Y-index were changed. Some institutions were increased such as National Cancer Institute, University of Cambridge, Yale University, University of California San Francisco, Massachusetts Institute of Technology, and University of California Berkeley and were decreased including Massachusetts General Hospital, University of Pennsylvania, University of Oxford, Johns Hopkins University, Brigham & Womens' Hospital, Columbia University, University of California San Diego, and University of Texas. Seven institutions in Table 5 with $\theta > 0.7854$ had more corresponding author publications for example Columbia University and Massachusetts Gen Hospital while nine institutions with $\theta < 0.7854$ had more first author publications for example University of California Los Angeles.

Everyone listed as an author of one article has made an independent material contribution to the manuscript (Coats 2009). Of the 3,022 top-cited publications with first and corresponding author information in SCI-Expanded, there were 22,335 authors from 62 countries. The percentages of publication with one, two, three, and four authors were 82, 13, 3.2, and 1.0 %, respectively. Only 1 % of 22,335 authors published at least five top-cited papers. Overall, top-cited publications (3,022) were published by 2,636 (12 %) first

Table 5 Characteristics of the top 20 institutions

Institute	TP	TP R (%)	SP R (%)	CP R (%)	FP R (%)	RP R (%)	θ	$j R (j)$
Harvard University, USA	231	1 (7.6)	1 (3.5)	1 (1.1)	1 (3.5)	1 (3.3)	0.7610	1 (145)
Stanford University, USA	120	2 (4.0)	3 (2.2)	2 (5.4)	3 (1.9)	2 (2.0)	0.8199	2 (82.1)
University of Texas, USA	108	3 (3.6)	4 (1.8)	3 (5.0)	4 (1.6)	4 (1.6)	0.7959	4 (67.2)
Massachusetts Institute of Technology (MIT), USA	106	4 (3.5)	5 (1.7)	4 (5.0)	2 (2.0)	3 (1.7)	0.7030	3 (77.3)
University of Washington, USA	97	5 (3.2)	8 (1.3)	5 (4.7)	5 (1.4)	5 (1.5)	0.7969	5 (61.5)
University of California San Diego, USA	87	6 (2.9)	6 (1.5)	7 (4.0)	7 (1.3)	8 (1.1)	0.7045	8 (52.5)
University of California Berkeley, USA	85	7 (2.8)	2 (2.2)	13 (3.3)	6 (1.4)	6 (1.4)	0.7854	6 (58.0)
Johns Hopkins University, USA	81	8 (2.7)	19 (0.82)	6 (4.2)	12 (0.86)	15 (0.73)	0.7023	14 (34.1)
University of California San Francisco, USA	79	9 (2.6)	6 (1.5)	10 (3.5)	8 (1.3)	7 (1.3)	0.7854	7 (53.7)
University of Michigan, USA	72	10 (2.4)	23 (0.74)	9 (3.7)	11 (1.0)	10 (1.0)	0.7854	10 (41.0)
Brigham & Womens' Hospital, USA	69	11 (2.3)	126 (0.15)	7 (4.0)	16 (0.73)	13 (0.79)	0.8288	15 (32.6)
Yale University, USA	66	12 (2.2)	23 (0.74)	12 (3.3)	9 (1.1)	9 (1.0)	0.7542	9 (45.3)
Massachusetts General Hospital, USA	61	13 (2.0)	126 (0.15)	10 (3.5)	56 (0.36)	31 (0.53)	0.9685	41 (19.4)
Washington University, USA	61	13 (2.0)	9 (1.3)	16 (2.6)	12 (0.86)	13 (0.79)	0.7454	13 (35.4)
University of Oxford, UK	59	15 (2.0)	29 (0.67)	14 (3.0)	21 (0.60)	27 (0.56)	0.7568	26 (24.8)
University California Los Angeles, USA	58	16 (1.9)	10 (1.1)	18 (2.6)	15 (0.79)	21 (0.66)	0.6947	16 (31.2)
National Cancer Institute, USA	57	17 (1.9)	10 (1.1)	20 (2.5)	10 (1.0)	12 (0.86)	0.7141	11 (39.7)
University of Cambridge, UK	53	18 (1.8)	14 (0.89)	21 (2.4)	14 (0.83)	11 (0.89)	0.8238	12 (36.8)
Columbia University, USA	50	19 (1.7)	37 (0.52)	18 (2.6)	33 (0.50)	15 (0.73)	0.9724	22 (26.6)
University of Pennsylvania, USA	50	19 (1.7)	46 (0.45)	16 (2.6)	27 (0.53)	31 (0.53)	0.7854	31 (22.6)

TP total publications, R rank, SP single institution publications, CP collaborative publications, FP first author publications, RP corresponding author publications, j Y-index parameter, publication intensity constant, θ Y-index parameter, publication character constant

authors, 2,481 (11 %) corresponding author. The top 12 productive authors who published more than 10 top-cited papers, were Lander E. S. (17 papers), Wang, J. (15), Collins, R. (14), Peto, R. (14), Akira, S. (13), Vogelstein, B. (13), Brown, P.O. (12), Wang, Y. (12), Yusuf, S. (11), Botstein, D. (11), Thun, M. J. (11), and Murray, T. (11).

Figure 3 shows distribution of the Y-index (j, θ) of the top 67 authors with $j > 4.00$. j is publication intensity constant, an author with a higher j indicates more papers as first or corresponding authors, and partake leadership role in more papers. Jemal, A. had 8 papers in which all were first author and corresponding author ($\theta = 0.7854$), and had the highest j of 11.3, followed by Lieber, C. M. ($j = 9.00$), Kresse, G. ($j = 8.49$), and Mirkin, C. A. ($j = 8.06$). θ , a publication character constant, differentiate its nature of leadership role. When $\theta > 0.7854$, means published more corresponding author papers and $\theta < 0.7854$, means published more first author papers. When $\theta = 0, j =$ number of first author articles and $\theta = \infty, j =$ number of corresponding author articles. Lieber, C. M. published 9 corresponding authors and none first author paper (9.00, ∞), followed by Jemal, A. and Mirkin, C. A. both had 8 corresponding author papers. Jemal, A. published 8 first author papers followed by Kresse, G. with 6. Alivisatos, A. P., Ridker, P. M., Botstein, D. and Dekker, C. had the same numbers of publications ($r = 5.00$). Botstein, D. and Dekker, C. ($\theta = 0.9273$) published more first author papers than Alivisatos, A. P. and Ridker, P. M. who published only five corresponding author papers but none first author paper ($\theta = \infty$). All authors in Fig. 3 had a $\theta \geq 0.7854$.

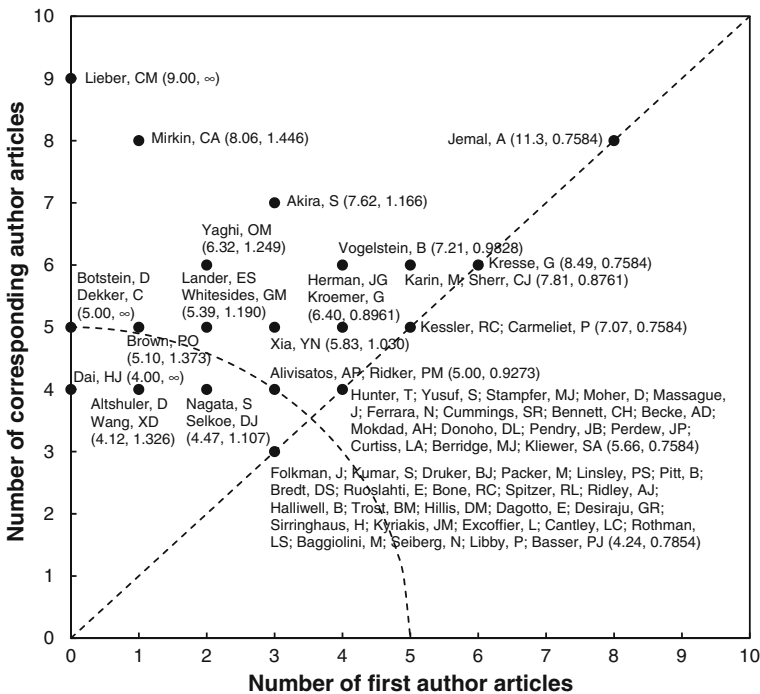


Fig. 3 Top 67 authors with Y-index ($i > 4.00$)

Conclusion

The 3,652 top-cited papers in nine document types published from 1991 to 2010 were analyzed. Of these 2,541 articles were published in 365 journals listed in 122 Web of Science categories. *Science*, *Nature*, *New England Journal of Medicine*, and *Cell* which ranked top four by top-cited articles, of the 8,073 journals in the JCR Science Edition in 2010, published most top-cited articles. The USA ranked first by all indicators, although the article with most citations was published by Thompson, Higgins, and Gibson from the European Molecular Biology Laboratory, Germany. Six of the top 12 articles with $TC_{2010} > 10,000$ were published as institution independent articles and nine articles were country independent from the USA, Germany, Canada, and Japan respectively. A new indicator, the Y-index was proposed and successfully applied to evaluate the publication character of authors, institutions, and countries. The number of first and corresponding author articles were similar for institutions and countries, but varied significantly among individuals.

Harvard University was the most productive institution, ranked top one in total, first author, and corresponding author top-cited publications, as indicated by the Y-index. Lander E. S. was the most productive author while Jemal A. had highest publication performance of first and corresponding author articles, as indicated by the Y-index. Both first author and corresponding author are major contributors to a published research work, but the quality of their contribution does differ significantly. It was shown that the Y-index can assist bibliometric researchers to look behind the usual indices of total publication, but also the character of contribution.

References

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J. H., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389–3402.
- Baltussen, A., & Kindler, C. H. (2004). Citation classics in anesthetic journals. *Anesthesia and Analgesia*, 98(2), 443–451.
- Becke, A. D. (1993). Density-functional thermochemistry. III. The role of exact exchange. *Journal of Chemical Physics*, 98(7), 5648–5652.
- Bennett, D. M., & Taylor, D. M. (2003). Unethical practices in authorship of scientific papers. *Emergency Medicine*, 15(3), 263–270.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., et al. (1998). *Crystallography and NMR system: A new software suite for macromolecular structure determination. Acta Crystallographica Section D-Biological Crystallography*, 54, 905–921.
- Brush, S. G. (1990). The most cited physical sciences publications in the 1945–1954 Science Citation Index. *Current Contents*, 20, 7–17.
- Burman, K. D. (1982). Hanging from the masthead—Reflections on authorship. *Annals of Internal Medicine*, 97(2), 602–605.
- Celayir, S., Sander, S., Ellicevik, M., Vural, A., & Celayir, A. C. (2008). The most commonly cited articles in pediatric surgical journals. *European Journal of Pediatric Surgery*, 18(3), 160–163.
- Chen, S. R., Chiu, W. T., & Ho, Y. S. (2005). Asthma in children: Mapping the literature by bibliometric analysis. *Revue Française d'Allergologie et d'Immunologie Clinique*, 45(6), 442–446.
- Chuang, K. Y., Wang, M. H., & Ho, Y. S. (2011). High-impact papers presented in the subject category of water resources in the essential science indicators database of the institute for scientific information. *Scientometrics*, 87(3), 551–562.
- Chiu, W.T., & Ho, Y.S. (2005). Bibliometric analysis of homeopathy research during the period of 1991 to 2003. *Scientometrics*, 63(1), 3–23.
- Coats, A. J. S. (2009). Ethical authorship and publishing. *International Journal of Cardiology*, 131(2), 149–150.

- Costas, R., & Bordons, M. (2011). Do age and professional rank influence the order of authorship in scientific publications? Some evidence from a micro-level perspective. *Scientometrics*, 88(1), 145–161.
- Dotson, B., & Slaughter, R. L. (2011). Prevalence of articles with honorary and ghost authors in three pharmacy journals. *American Journal of Health-System Pharmacy*, 68(18), 1730–1734.
- Egghe, L. (2006). Theory and practise of the g-index. *Scientometrics*, 69(1), 131–152.
- Engers, M., Gans, J. S., Grant, S., & King, S. P. (1999). First-author conditions. *Journal of Political Economy*, 107(4), 859–883.
- Fu, H. Z., Ho, Y. S., Sui, Y. M., & Li, Z. S. (2010). A bibliometric analysis of solid waste research during the period 1993–2008. *Waste Management*, 30(12), 2410–2417.
- Furlan, J. C., & Fehlings, M. G. (2006). A web-based systematic review on traumatic spinal cord injury comparing the “citation classics” with the consumers’ perspectives. *Journal of Neurotrauma*, 23(2), 156–169.
- Gaeta, T. J. (1999). Authorship: “Law” and order. *Academic Emergency Medicine*, 6(4), 297–301.
- Garfield, E. (1955). Citation indexes to science: A new dimension in documentation through the association of ideas. *Science*, 122(3159), 108–111.
- Garfield, E. (1964). “Science Citation Index”—New dimension in indexing. *Science*, 144(3619), 649–654.
- Garfield, E. (1970). Citation indexing for studying science. *Nature*, 227(5259), 669–671.
- Garfield, E. (1989). An introduction to the 102 most cited life sciences publications in the new 1945–1954 Science Citation Index. 1. Titles, journals, and research fronts by Dixon, Bernard. *Current Contents*, 15, 3–10.
- Garfield, E. (1990). The Russians are coming. 2. The top 50 soviet papers most cited in the 1973–1988 Science Citation Index and a look at 1988 research fronts. *Current Contents*, 25, 3–13.
- Garfield, E., & Welljamsdorff, A. (1992). Of Nobel class: A citation perspective on high-impact research authors. *Theoretical Medicine and Bioethics*, 13(2), 117–135.
- Gehanno, J. F., Takahashi, K., Darmoni, S., & Weber, J. (2007). Citation classics in occupational medicine journals. *Scandinavian Journal of Work, Environment and Health*, 33(4), 245–251.
- Han, J. S., & Ho, Y. S. (2011). Global trends and performances of acupuncture research. *Neuroscience and Biobehavioral Reviews*, 35(3), 680–687.
- Hennessey, K., Afshar, K., & MacNeily, A. E. (2009). The top 100 cited articles in urology. *Canadian Urological Association Journal*, 3(4), 293–302.
- Herbert, H., & Müller-Hill, B. (1995). Quality and efficiency of basic research in molecular biology: A bibliometric analysis of thirteen excellent research institutes. *Research Policy*, 24(6), 959–979.
- 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(46), 16569–16572.
- Ho, Y. S., Satoh, H., & Lin, S. Y. (2010). Japanese lung cancer research trends and performance in Science Citation Index. *Internal Medicine*, 49(20), 2219–2228.
- Hsieh, W. H., Chiu, W. T., Lee, Y. S., & Ho, Y. S. (2004). Bibliometric analysis of patent ductus arteriosus treatments. *Scientometrics*, 60(2), 205–215.
- Huth, E. J. (1986). Guidelines on authorship of medical papers. *Annals of Internal Medicine*, 104(2), 269–274.
- Iijima, S. (1991). Helical microtubules of graphitic carbon. *Nature*, 354(6348), 56–58.
- Jemal, A., Siegel, R., Ward, E., Hao, Y. P., Xu, J. Q., & Thun, M. J. (2009). Cancer Statistics, 2009. *CA-A Cancer Journal for Clinicians*, 59(4), 225–249.
- Jin, B. H. (2006). h-index: An evaluation indicator proposed by scientist. *Science Focus*, 1(1), 8–9.
- Jin, B. H., Liang, L. M., Rousseau, R., & Egghe, L. (2007). The R- and AR-indices: Complementing the h-index. *Chinese Science Bulletin*, 52(6), 855–863.
- Jones, T. A., Zou, J. Y., Cowan, S. W., & Kjeldgaard, M. (1991). Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Crystallographica Section A*, 47(1), 110–119.
- Lee, M., & Park, H. W. (2012). Exploring the web visibility of world-class universities. *Scientometrics*, 90(1), 201–218.
- Lefavre, K. A., Shadgan, B., & O’Brien, P. J. (2011). 100 most cited articles in orthopaedic surgery. *Clinical Orthopaedics and Related Research*, 469(5), 1487–1497.
- Li, L. L., Ding, G. H., Feng, N., Wang, M. H., & Ho, Y. S. (2009a). Global stem cell research trend: Bibliometric analysis as a tool for mapping of trends from 1991 to 2006. *Scientometrics*, 80(1), 39–58.
- Li, Z., & Ho, Y. S. (2008). Use of citation per publication as an indicator to evaluate contingent valuation research. *Scientometrics*, 75(1), 97–110.
- Li, J. F., Zhang, Y. H., Wang, X. S., & Ho, Y. S. (2009b). Bibliometric analysis of atmospheric simulation trends in meteorology and atmospheric science journals. *Croatica Chemica Acta*, 82(3), 695–705.

- Lin, H. W., Yu, T. C., & Ho, Y. S. (2011). A systemic review of the human papillomavirus studies: Global publication comparison and research trend analyses from 1993 to 2008. *European Journal of Gynaecological Oncology*, 32(2), 133–140.
- Li, J.F., Wang, M.H., & Ho, Y.S. (2011). Trends in research on global climate change: A Science Citation Index Expanded-based analysis. *Global and Planetary Change*, 77(1–2), 13–20.
- Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods*, 25(4), 402–408.
- Malin, M. V. (1968). Science Citation Index—A new concept in indexing. *Library Trends*, 16(3), 374–387.
- Mattsson, P., Sundberg, C. J., & Laget, P. (2011). Is correspondence reflected in the author position? A bibliometric study of the relation between corresponding author and byline position. *Scientometrics*, 87(1), 99–105.
- Nieri, M., Saletta, D., Guidi, L., Buti, J., Franceschi, D., Mauro, S., et al. (2007). Citation classics in periodontology: A controlled study. *Journal of Clinical Periodontology*, 34(4), 349–358.
- Ohba, N., & Nakao, K. (2010). The 101 most frequently cited articles in ophthalmology journals from 1850 to 1949. *Archives of Ophthalmology*, 128(12), 1610–1617.
- Ohba, N., Nakao, K., Isashiki, Y., & Ohba, A. (2007). The 100 most frequently cited articles in ophthalmology journals. *Archives of Ophthalmology*, 125(7), 952–960.
- Ollerton, J. E., & Sugrue, M. (2005). Citation classics in trauma. *Journal of Trauma-Injury Infection and Critical Care*, 58(2), 364–369.
- Paladugu, R., Schein, M., Gardezi, S., & Wise, L. (2002). One hundred citation classics in general surgical journals. *World Journal of Surgery*, 26(9), 1099–1105.
- Perdew, J. P., Burke, K., & Ernzerhof, M. (1996). Generalized gradient approximation made simple. *Physical Review Letters*, 77(18), 3865–3868.
- Posada, D., & Crandall, K. A. (1998). MODELTEST: Testing the model of DNA substitution. *Bioinformatics*, 14(9), 817–818.
- Riesenberg, D., & Lundberg, G. D. (1990). The order of authorship—Who's on first. *JAMA-Journal of the American Medical Association*, 264(14), 1857.
- Rosenberg, A. L., Tripathi, R. S., & Blum, J. (2010). The most influential articles in critical care medicine. *Journal of Critical Care*, 25(1), 157–170.
- Schein, M., Paladugu, R., & Sutija, V. G. (2000). What American surgeons read: a survey of a thousand Fellows of the American College of Surgeons. *Current Surgery*, 57(3), 252–258.
- Sheldrick, G. M. (2008). A short history of SHELX. *Acta Crystallographica Section A*, 64(1), 112–122.
- Singh, S. (2009). Criteria for authorship. *Indian Journal of Dermatology Venereology and Leprology*, 75(2), 211–213.
- Slone, R. M. (1996). Coauthors' contributions to major papers published in the AJR: Frequency of undeserved coauthorship. *American Journal of Roentgenology*, 167(3), 571–579.
- Spek, A. L. (2009). Structure validation in chemical crystallography. *Acta Crystallographica Section D-Biological Crystallography*, 65, 148–155.
- Suk, F. M., Lien, G. S., Yu, T. C., & Ho, Y. S. (2011). Global trends in *Helicobacter pylori* research from 1991 to 2008 analyzed with the Science Citation Index Expanded. *European Journal of Gastroenterology and Hepatology*, 23(4), 295–301.
- Tanaka, H., & Ho, Y. S. (2011). Global trends and performances of desalination research. *Desalination and Water Treatment*, 25(1–3), 1–12.
- The role and responsibilities of coauthors. (2008). *Lancet*, 372(9641), 778.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 25(24), 4876–4882.
- Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22(22), 4673–4680.
- Wang, M. H., Li, J. F., & Ho, Y. S. (2011). Research articles published in water resources journals: A bibliometric analysis. *Desalination and Water Treatment*, 28(1–3), 353–365.
- Wang, M. H., Yu, T. C., & Ho, Y. S. (2010). A bibliometric analysis of the performance of Water Research. *Scientometrics*, 84(3), 813–820.
- Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-Item short-form health survey (SF-36). I. Conceptual-framework and item selection. *Medical Care*, 30(6), 473–483.
- Zuckerman, H. A. (1968). Patterns of name ordering among authors of scientific papers: A study of social symbolism and its ambiguity. *American Journal of Sociology*, 74(3), 276–291.