University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal) | Libraries at University of Nebraska-Lincoln

March 2021

Scientometric Analysis of Public Health Literature: A Study based on Scopus Database

Dr. Md. Nazmul Islam
University of Rajshahi, Bangladesh, nazmul.islam.81@gmail.com

Dr. Md. Shariful Islam
University of Rajshahi, Bangladesh, sharif6islam@yahoo.com

Dr. Partha Biplob Roy
University of Rajshahi, Bangladesh, pbroy2000@yahoo.co.in

Scientometric Analysis of Public Health Literature: A Study based on Scopus Database

Dr. Md. Nazmul Islam
Associate Professor, Department of Information Science and Library Management, University of Rajshahi, Bangladesh, E-mail: nazmul.islam.81@gmail.com

Dr. Md. Shariful Islam
Corresponding author and Professor, Department of Information Science and Library Management, University of Rajshahi, Bangladesh, E-mail: sharif6islam@yahoo.com

Dr. Partha Biplob Roy
Professor, Department of Information Science and Library Management, University of Rajshahi, Bangladesh, E-mail: pbroy2000@yahoo.co.in

Abstract

Purpose – The main purpose of this study is to evaluate the progress and development of the global literature on public health and review several components of scientometric study including the patterns of growth of literature, authorship, author collaboration, and productivity.

Design/methodology/approach – The current study is exploratory research in nature reviewing secondary literature extracted from a Scopus database and analytical with the application of suitable statistical and scientometric tools to strengthen the empirical validity. Public health literature was assessed quantitatively using scientometric indicators from 2000 to 2015.

Findings – The year 2003 was the most productive and 2015 was the least productive year in Annual Growth Ratio (AGR) of public health published literature across the globe during the period under study. The average Relative Growth Rate (RGR) globally was 0.23 and at this rate, the literature of public health doubles every 4.16 years. Globally, the average Collaborative Coefficient (CC) value for global authors was 0.37 demonstrating that there was no significant magnitude of collaboration among worldwide authors. The average Productivity Per Author (PPA) for global authors is 0.49 which means public health authors produced less than half a publication each year during the study period.

Originality/value – The paper is one of the very few studies which focuses on Scientometric analysis of public health literature using Scopus database. The present research gives a direction to determine the level of collaboration by a Collaborative Coefficient (CC) value which was never been calculated previously.

Keywords – Scientometric study, Bibliometric study, Public health (PH), Author Collaboration, Author Productivity, Scopus database.

Paper type Research paper

1. Introduction

It has become increasingly important for library professionals/subject specialists to know the nature of research publications and pattern of growth of research output of different disciplines, and over different periods, due to the ever-expanding knowledge resources available in different forms like books, periodical articles, reports, theses, patents, proceedings, web pages etc. Early 20th century statistical bibliography was of assistance to this challenge, and during the latter part of that century was called Bibliometrics or Informetrics or Webometrics or Scientometrics, often abbreviated to BIWS. With these techniques it is easily possible to trace out many facets of research, including - the priority areas, thriving fields, future growth of research output, values of research works, age of literature used and the information needs of researchers, scientists and subject experts, importance of different types of publication, the shape of development of a discipline at different times, regions, affiliated institutions and subjects, and those which have helped in policy/decision making as well as scientific communication.

The two terms, 'bibliometrics' and 'scientometrics', were concurrently introduced almost by Pritchard and by Nalimov and Mulchenko in 1969 (Glänzel, n.d.). Pritchard treated 'Bibliometric' as *Statistical Bibliography,* denoting the application of mathematics and statistical methods to books and other media of communication (Pritchard, 1969). Nalimov and Mulchenko used the Russian word 'Naukometriya' to explain scientometrics as *"the application of quantitative methods to analyze science as an information process"* (Nalimov & Mulchenko, 1969). Therefore scientometrics is limited to the measurement of science communication, whereas bibliometrics is intended to deal with basic information processes although now-a-days both terms are used interchangeably (Glänzel, n.d.). The term scientometrics became more well-known, however, once the journal "Scientometrics" first appeared in 1978 (Garfield, 2009). The term 'Informetrics' comes from German term 'informetrie', and was first used by Nacke in 1979, and was treated as a broader term for both bibliometrics and scientometrics leaning to policy studies (Hood & Wilson, 2001; Brookes, 1990).

Public health, on the other hand, denotes to those activities which are used to prevent diseases, ensure sound health, and extend life for the people of the society. Therefore, the main objective of public health is to assure conditions for people to be healthy. In two different studies conducted by Jakovljevic & Ogura and Porter found the dynamic history of public health, which has been formed by the evolution of diseases. The study of health research has been shifted from the sanitary developments and the control of communicable diseases to the impact of wide spreading and infectious diseases, and the involvement of social movement as the controller of rampantcalamities.

Therefore, public health research covers a wide range of disciplines including health economics and socio-political aspects of health, etc. (As cited in Merigo & Nunez, 2016).

It is necessary to examine the status of public health research globally, its growth rate, authorship pattern, collaborative nature, and author productivity. It was observed that no such study has been conducted at globally on the development pattern of public health literature although a few numbers of journal articles, conference proceedings, and reports were published on the scientometric study of public health at a micro (institutions) or macro(country) level only.

2. Review of related research

The present era of information explosion demands more research and development is in not only natural science subject but also other subject fields. To get expected expertise in their field of specialization and fulfill the knowledge gap, the researchers of today's world are increasingly interested to work in collaboration. According to Aristotle's metaphysics theory, the whole is more than some of its parts. This means that combining forces produces not the only better product but also maximum product. The main logic behind this theory is that successful integration normally produces a synergistic effect and greater total impact than if each author works separately.

The teamwork what we called in scientometric study as collaborated works among research scholars has great importance in the development and dissemination of subject. The synergistic efforts to work with a subject might open new door of thinking which is impossible in few cases to solve the problem individually. The investigation of authorship pattern, author collaboration, and author productivity on public health publications can reflect the differentaspects of the subject (Ding, Foo & Chowdhury, 1999).

Several attempts have already been made all over the world to measure authors' productivity, authorship patterns and author collaboration (Hemala & Kavitha, 2016; Gajbe & Sonawane, 2015; Kumar& Naqvi, 2014; Rakhi, 2014; Jimenez-Fanjul, Maz-Machado & Bracho-Lopez, 2013; Heidari and Safavi, 2013; Thilakar and Ponnudurai, 2013; Arya & Sharma, 2012; Elango & Rajendran, 2012; Pillai, 2007; Yazit & Zainab, 2007; Udofia, 2002; Ding, Foo & Chowdhury, 1999). Several formulas and indicators have been devised to study the average number of authors per paper, the collaborative pattern of authorship on a subject, the proportion of single and multi-authored papers, etc. Some important formulas in this regard are Collaborative Index (CI) devised by Lawani in 1980, Degree of Collaboration (DC) by Subramanyam in 1983, Collaborative Coefficient (CC) by Ajiferuke, Burrell & Tague in 1988, etc.

Measurement of growth and development of any subject is a common phenomenon all over the world. Several studies have been conducted globally on the application of scientometric techniques in various subject fields. Some of such subject fields where techniques of scientometric have already been applied to measure the quantum of growth and development as Metallurgy and Material Sciences (Sandha, 2001); Materials Science and Engineering (Rao, 2005); Software (Sahoo, 2006); Chemical Sciences (Meera, 2007); Building materials (Senapati, 2009); Agriculture (Ravanan, 2012); Physics (Sedam, 2013); Epidemiology (Mahendran, 2014); Fashion Technology (Manimegalal, 2014); Textile Technology (Packiyaraj, 2014); Nano thin films (Prabakar, 2014); Genetic Engineering (Balasubramani, 2015); Brain tumor (Ramesh, 2015); Wireless communication (Manickaraj, 2015); Malaria research (Meena, 2015); Human DNA (Murugiah, 2015); Mems (Narayanan A L, 2015); Swine Influenza (Sivakami, 2015); Nuclear power generation (Venkatesan, 2015); Biotechnology (Tejashwini, 2016); Rabies (Sachithanantham, 2017).

On the other hand, very few publications were found on the scientometric study of public health. Three of such works have been found at macro level (country's production on public health) (Kalita, Shinde & Patel, 2015; Donner, Chi & Aman, 2014; Macias-Chapula *et al.*, 2008) and two works on continent's production on public health (Chuang *et al.*, 2011; Clarke, *et al.*, 2007) but no work was found assessing public health literature globally.

3. Research objectives

The study was designed generally to assess public health literature indexed in the Scopus database during the period 2000-2015 in terms of growth pattern, author collaboration and productivity. To achieve this, the following special objectives were devised:

i. To investigate growth and development of public health literature using several indicators;
ii. To calculate authorship patterns and author collaboration using various indices;
iii. To measure authors' productivity using various scientometric indicators.

4. Research Methodology

The present study is primarily exploratory reviewing public health-related scholarly output extracted from Scopus database from 2000 to 2015. It can also be considered scientometric research that helps to make decisions based on scholarly communication. To extract data from the Scopus database the search terms "public health" and period "2000-2015" were used: "(TITLE-ABS-KEY (Public Health) AND PUBYEAR > 1999 AND PUBYEAR < 2016)". During the search period, all types of documents relating to public health literature, including research articles, reviews, books, conference proceedings, etc. were identified. A total of 3,72,260 records for the scientometric research were downloaded from the Scopus database on 23 November 2016. Various **statistical tools** such as arithmetic mean, percentage, cumulative percentage, average, time series analysis, simple linear regression, t-test, etc. have been used for the study.The t-test has been carried out at 0.01 level of significance. Various **quantitative indicators** used in scientometric study for analyzing growth and development of literature such as Annual Growth Ratio (AGR), Average Annual Growth Ratio (AAGR), Percent Growth Rate (PR), Compound Annual Growth Rate (CAGR), Relative Growth Rate (RGR), Doubling time (Dt), Future Growth Rate (FGR) have been applied. For analyzing the collaborative pattern and author productivity of literature quantitative indicators like Collaborative Index (CI), Degree of Collaboration (DC), Collaborative Coefficient (CC), Revised Collaborative Coefficient (RCC), Average Author Per Paper (AAPP), Productivity Per Author (PPA) were used. The data on public health literature extracted from Scopus database was analyzed using MS-Excel and SPSS (version-24.0) software.

5. Data Analysis and Findings

5.1 Assessment of growth of literature

The quantum of Public Health (PH) literature can be analyzed differently to measure the growth of literature using different scientometric indicators and techniques.

5.1.1 Annual Growth Ratio (AGR) and Average Annual Growth Ratio (AAGR) of PH literature

During the current study period from 2000 to 2015, there were 3,72,260 publications on public health enlisted in the Scopus database. Table 1 shows the year-wise growth of public health literature during the study period.

Table 1: Annual Growth Ratio (AGR) and Average Annual Growth Ratio (AAGR) of PH literature

Year	Four Year Grouping	Publications	Publications	AGR	AAGR
2000	2000-2003	11,594	56,420	1: 1.06	1.11
2001		13,325		1: 1.15	
2002		14,683		1: 1.10	
2003		16,818		1: 1.15	
2004	2004-2007	18,329	80,594	1: 1.09	1.07
2005		19,668		1: 1.07	
2006		20,845		1: 1.06	
2007		21,752		1: 1.04	
2008	2008-2011	22,750	1,01,628	1: 1.05	1.07
2009		24,289		1: 1.07	
2010		26,222		1: 1.08	
2011		28,367		1: 1.08	
2012	2012-2015	31,044	1,33,618	1: 1.09	1.05
2013		32,297		1: 1.04	
2014		35,430		1: 1.10	

Year	Four Year Grouping	Publications	Publications	AGR	AAGR
2015		34,847		1: 0.98	
Total		3,72,260	3,72,260		

Note: There were 10918 publications in 1999 (Source: Scopus); Cells in highlighted font shows highest and lowest values.

Annual Growth Ratio (AGR) is calculated as the present number of publications divided by the previous number of publications. The Annual Growth Ratio (AGR) of PH literature annually varies from 0.98 to 1.15. The most productive years recorded were 2001 and 2003 (1: 1.15), and the year 2015 had the lowest AGR (1:0.98) of PH literature. **Average Annual Growth Ratio (AAGR)** is calculated as the summation of the values of the specific period of interval divided by the number of period interval. The study period (2000-2015) was grouped into four class of intervals each of which represents 4 years. The period from 2000 to 2003 had the highest AAGR and the lowest AAGR was observed during 2012-2015. The highest number of research was conducted during the period from 2012 to 2015 (35.89%) followed by a period of 2008 to 2011 (27.30%).

To analyze the relationship between years' increasing and growth of literature the t-test has been carried out to see whether or not the regression coefficient is significant. In this test, year values are independent variables and publication values are dependent variables.

Table 2: Regression Coefficients

Regression Statistics			***Coefficients***	***Standard Error***	***t Stat***	***P-value****
Multiple R	0.993846053	Intercept	-3142425.449	94299.4837	-33.3239	9.77E-15
R Square	0.987729977	Year	1576.932353	46.97346704	33.5707	8.82E-15
Adjusted R Square	0.986853547					
Standard Error	866.1479354					
Observations	16					

Note: *Significant at p<0.01

Table 2 shows that the value of R square is 0.9877. That means 98% of the variability in the publication is explained by the regression line or by the regression of year on publication. The result of the "t-test" showed that there is a significant relationship between the progressing of years and the growth of literature ($p<0.01$).

5.1.2 Percent Growth Rate (PR), Compound Annual Growth Rate (CAGR), Relative Growth Rate (RGR) and Doubling time (Dt) of PH literature

Percent (Straight-Line) Growth Rate (PR) is calculated as subtraction between the number of publications of the present yearand the number of publications of the previous yeardivided by the number of publication in previous year**.** The formula of PR (Parker, 2002) is:

$$PR= \frac{Presentyearvalue - Previous\ year\ value}{Previousyearvalue} x\ 100 \quad \textbf{[Eq. 1]}$$

Compound Annual Growth Rate (CAGR) is calculated as the number of publications of the present year divided by the number publication previous year to the power of one divided by the period length and subtracts one from the subsequent result (Murphy, 2019).The formula of CAGR (Kumar and Kaliyaperumal, 2015) is:

$$CAGR=\left(\frac{EndingValue}{BeginningValue}\right)^{\left(\frac{1}{Numberofyears}\right)}-1 \quad \textbf{[Eq. 2]}$$

The concept of **Relative Growth Rate (RGR)** has been derived from botany to express growth in terms of a rate of increase in size per unit of size. V.H. Blackman called it 'efficiency index' (Hunt, 1990) which is in later used to measure relative growth of literature for a specific period. RGR can be calculated through the following equation (Hunt, 1982):

$$\text{RGR or R}=\frac{Log_e W_2 - Log_e W_1}{T_2 - T_1} \quad \textbf{[Eq. 3]}$$

Whereas,

RGR or R = mean relative growth rate over the specific period of interval

$Log_e W_1$ = log of the initial number of articles/pages

$Log_e W_2$ = log of the final number of articles/pages after a specific period of interval

T_2-T_1 = Unit difference between initial time and final time.

Dt (Doubling time) is directly related to RGR and is defined as the time required for the articles to become two-fold of the existing amount. The natural logarithm has been used here by taking the value of $log_e 2$ as 0.693 (Keshava, 2004). Therefore, the corresponding doubling time for articles can be calculated by the following equation (Kumar and Kaliyaperumal, 2015):

Dt(a[1]) $=\frac{\ln(2)}{RGR}$ **or** $\frac{0.693}{RGR}$ **[Eq. 4]**

Table 3: PR, CAGR, RGR and Dt(a) of PH literature

Year	Records	Difference between two year	Cum	PR	CAGR	W_1	W_2	RGR	Dt(a)
2000	11,594	676	11,594	6%	0.06	-	9.36	-	-
2001	13,325	1731	24,919	15%	0.15	9.36	10.12	0.76	0.91
2002	14,683	1358	39,602	10%	0.10	10.12	10.59	0.47	1.47
2003	16,818	2135	56,420	15%	0.15	10.59	10.94	0.35	1.98
2004	18,329	1511	74,749	9%	0.09	10.94	11.22	0.28	2.47
2005	19,668	1339	94,417	7%	0.07	11.22	11.46	0.24	2.89
2006	20,845	1177	1,15,262	6%	0.06	11.46	11.65	0.19	3.65
2007	21,752	907	1,37,014	4%	0.04	11.65	11.83	0.18	3.85

[1] Here Doubling time (Dt) has been calculated based on articles.

Year	Records	Difference between two year	Cum	PR	CAGR	W_1	W_2	RGR	Dt(a)
2008	22,750	998	1,59,764	5%	0.05	11.83	11.98	0.15	4.62
2009	24,289	1539	1,84,053	7%	0.07	11.98	12.12	0.14	4.95
2010	26,222	1933	2,10,275	8%	0.08	12.12	12.26	0.14	4.95
2011	28,367	2145	2,38,642	8%	0.08	12.26	12.38	0.12	5.77
2012	31,044	2677	2,69,686	9%	0.09	12.38	12.51	0.13	5.33
2013	32,297	1253	3,01,983	4%	0.04	12.51	12.62	0.11	6.30
2014	35,430	3133	3,37,413	10%	0.10	12.62	12.73	0.11	6.30
2015	34,847	-583	3,72,260	-2%	-0.02	12.73	12.83	0.10	6.93
Total	3,72,260	Average PR& CAGR		8%	0.08	Average RGR & Dt(a)		0.23	4.16

Note: There were 10918 publications in 1999 (Source: Scopus); Cells in highlighted font shows highest and lowest values.

Table 3 shows a steady growth of literature from 2000 to 2014, the differences in the number of publications listed in each year remains 676 to 3,133 during this period. The most productive year in terms of increasing publication than the previous year is 2014 (3,133 publications in this year, 9.52% in total). Only in the year 2015, the number of publications decreased if we compare with the literature published in the previous year (-583). There is an obvious variation on the growth of literature during the study period (2000-2015). During that period, PR varies from -2% to 15% and average PR is 8%. CAGR values also vary during the study period (-0.02 to 0.15). The lowest CAGR was during 2015 (-0.02) and the highest CAGR during 2001 and 2003. The RGR values in the field of public health during the period of 2000 to 2015 lies between 0.10 and 0.76. A downward trend for RGR values has been observed during the period 2000-2015. The Dt(a) values range between 0.91 to 6.93 and average Dt(a) value is 4.16. This means that the literature published in public health doubles in every 4.16 year in the period of the study.

5.1.3 Future Growth Rate (FGR) of PH literature

Using simple linear trends method of the period under study, future growth of expected literature on public health can be estimated (Table 4). Straight line equation can be applied for future growth by time series analysis. Straight line equation: Y_e=a+bX. Since Σx=0 a= Σ Y/N= 372260/16 = 23266.25 b= ΣXY/Σx2 = 536157/340= 1576.93235. As per straight line equation: Y_e=a+bX. Estimated literature in 2017 will be 38247.10733. Where X= 2017-2007.5 = 9.5 a = 23266.25 b = 1576.93235. So Y_e=a+bX = 23266.25 + 1576.93235*9.5= 38247.10733.

Table 4: Simple linear method for the future growth of PH literature

Serial	Year (x)	Publications (y)	X	X^2	Xy	Year	Expected publications
1.	2000	11,594	-7.5	56.25	-86955	2017	38,247.11
2.	2001	13,325	-6.5	42.25	-86612.5	2018	39,824.04
3.	2002	14,683	-5.5	30.25	-80756.5	2019	41,400.97

4.	2003	16,818	-4.5	20.25	-75681	2020	42,977.9
5.	2004	18,329	-3.5	12.25	-64151.5	2021	44,554.84
6.	2005	19,668	-2.5	6.25	-49170	2022	46,131.77
7.	2006	20,845	-1.5	2.25	-31267.5	2023	47,708.7
8.	2007	21,752	-0.5	0.25	-10876	2024	49,285.63
9.	2008	22,750	0.5	0.25	11375	2025	50,862.57
10	2009	24,289	1.5	2.25	36433.5	2026	52,439.5
11.	2010	26,222	2.5	6.25	65555	2027	54,016.43
12.	2011	28,367	3.5	12.25	99284.5	2028	55,593.36
13.	2012	31,044	4.5	20.25	139698	2029	57,170.3
14.	2013	32,297	5.5	30.25	177633.5	2030	58,747.23
15.	2014	35,430	6.5	42.25	230295	2031	60,324.16
16.	2015	34,847	7.5	56.25	261352.5	2032	61,901.09
Total	32120	3,72,260		340	536157		
Average	2007.5	23,266.25					

The future growth of literature on public health can be predicted using linear trends method. The future growth of literature on public health from the period of 2017-2032 has been calculated using the base year 2015 and shown in Table 4. The result shows an expected positive increase of literature each year. That means an increasing trend of literature on public health might be observed from the year 2017 to 2032 based on the RGR of 2000-2015. The search was conducted on November 2016. Based on that data, future growth of literature for the following year was calculated.

Table 5: Comparison of expected publications with the exact number of publication

Year	Expected number of publications	Exact number of publication	Difference
2017	38,247.11	40,993	+2,745.89
2018	39,824.04	44,812	+4,987.96

We can now compare our expected value with the exact number of publication on public health literature shown in table 5.In 2017, the Scopus database indexed 40,993 publication on public health literature, which is higher (+2,745.89) than our expected publication (38,247.11) calculated in 2016. In 2018 the difference between the expected number of publication with the exact number of publication was +4,987.96.

The t-test was carried out due to the quantitative nature of two dependent variables (existing and future growth of literature) and the qualitative nature of the independent variable (year) to see the mean difference of two dependent variables for significance.

Table 6: t-Test: Paired Two Sample for Means

	Variable 1	*Variable 2*
Mean	50074.09995	23266.25
Variance	56365554.43	57065752.73

Observations	16	16
Pearson Correlation	0.993846053	
Hypothesized Mean Difference	0	
Df	15	
t Stat	128.1479056	
P(T<=t) one-tail	1.61378E-24	
t Critical one-tail	1.753050325	
P(T<=t) two-tail	3.22756E-24*	
t Critical two-tail	2.131449536	

Note: *significant at p<0.01

The mean of the two dependent variables is 50074.09 for the variables of future growth (2017-2032) and 23266.25 for the variables of existing growth (2000-2015). Table 6 depicts that p-value is much lower than 0.01 which indicates there is a strong positive mean relationship between the existing and future growth of literature.

5.2 Identification of authorship pattern, author collaboration, and author productivity

5.2.1 Authorship pattern

Table 7 shows that the majority of publications under survey were published by collaborative authorship (53.34%). While a good number of literature were published in public health subject by single authorship (44.05%).

Table 7: Year-wise authorship pattern on PH literature

Year	Anon	Single Authors	Two Authors	Three Authors	Three + Authors	Total
2000	518	4,766	1,980	1,011	3,319	11,594
2001	684	5,319	2,613	1,703	3,006	13,325
2002	828	6,745	2,829	1,900	2,381	14,683
2003	960	7,889	3,001	1,598	3,370	16,818
2004	1003	8,025	2,875	1,922	4,504	18,329
2005	968	8,596	3,568	1,836	4,700	19,668
2006	674	9,149	3,478	2,565	4,979	20,845
2007	498	11,578	4,571	1,989	3,116	21,752
2008	425	10,583	4,002	2,205	5,535	22,750
2009	420	12,421	3,181	2,969	5,298	24,289
2010	455	12,898	5,735	2,627	4,507	26,222
2011	396	11,556	7,895	3,005	5,515	28,367
2012	469	10,527	8,698	3,102	8,248	31,044
2013	445	13,526	7,485	3,589	7,252	32,297
2014	517	15,520	8,756	3,901	6,736	35,430
2015	455	14,874	8,575	3,047	7,896	34,847
Total	9715 (2.61%)	1,63,972 (44.05%)	79,242 (21.29%)	38,969 (10.47%)	80,362 (21.59%)	3,72,260

5.2.2 Author collaboration

Lawani (1980) devised the **Collaborative Index** (CI) to measure the mean number of authors per paper. The formula of CI is:

$$CI= \frac{\sum_{j=1}^{k} jfj}{N} \quad \textbf{[Eq. 5]}$$

Where,

j = types of joint or collaborated author *i.e.* single author, two authors, three authors, etc.

fj = frequency of joint or collaborated author *i.e.* under joint/collaborated authors how many numbers of the research paper published on a subject during a certain period

N= Total number of the research paper published on a subject during a certain period

K= Greatest number of authors per paper on a subject.

CI can be calculated as the total number of authors divided by the total number of research articles published on a certain subject during a certain period. CI has some acute disadvantages also, such as single-authored paper has no collaboration, but it gives non-zero weight (of 1) to them, and it has no upper limit *i.e.* the value of CI neither lies between 0 and 1, and it is not expressible in terms of percentage.

Subramanyam (1983) proposed **the Degree of Collaboration** (DC) to measure the proportion of multi-authored papers. DC denotes the ratio of collaborated work to the total number of works of a subject field.

He devised the following formula to define the degree of collaboration:

$$DC = 1-\frac{fi}{N} \quad \textbf{[Eq. 6]}$$

Where DC means Degree of Collaboration

fi = single-authored papers;

N= Total number of the publication.

DC can be interpreted as a degree, *i.e.*, it lies between 0 and 1. A value of 1 means maximum collaboration. It always ranks higher in a discipline with a higher number of multi-authored papers though DC does not differentiate among levels of multiple authorships (Kumar & Naqvi, 2014; Ajiferuke, Burell & Tague, 1988).

Researchers in this area noted that the two collaborative measures *i.e.* both CI and DC had some inadequacies, which were removed by incorporating the merits of both, and devised a new measure by Ajiferuke and his team in 1988 called **Collaborative Coefficient** (CC). The value of CC can be calculated by the following formula (Ajiferuke, Burell & Tague, 1988):

$$CC = 1- \frac{\sum_{j=1}^{k}(\frac{1}{j})fj}{N} \quad \textbf{[Eq. 7]}$$

Where CC= Collaborative Coefficient

Fj = Number of authored papers in a subject during a certain period

N = Total number of research published in a subject during a certain period

K= the greatest number of authors per papers

The value of CC lies between zero and one. The value zero is corresponding to single authorship and whatever number is closer to one indicates more collaboration between authors. The value of CC can exactly lay "o" if all the frequencies remain under single authorship. Nevertheless, for maximal collaboration, CC fails to yield exactly one. That means the value of CC does not produce one in the case of all authors who are contributing as co-authors in the publication. To overcome from this situation some modifications has been done on CC called "**Revised Collaborative Coefficient** (RCC)" by Egghe and also called "Modified Collaborative Coefficient (MCC)" by Savanur and Srikanth. It is the normalized version of CC and is defined as the following formula (Todeschini & Baccini, 2016; Savanur & Srikanth, 2010; Egghe, 1991):

$$RCC = \frac{N}{N-1} \cdot \left(1 - \frac{\sum_{j=1}^{k}\left(\frac{1}{j}\right) f j}{N}\right) \quad \textbf{[Eq. 8]}$$

Table 8 shows the values of CI, DC, CC, and RCC. The CI is ranging from 2.26 (2012) to 1.80 (2007) with an average of 2.04 per paper which implies that the research team of just above two is typical in the field of public health.

Table 8: CI, DC, CC, and RCC of PH authors according to year

Year	CI	SA	TL	SA/TL	DC	CC	$\frac{N}{N-1}$	RCC
2000	2.16	4,766	11,594	0.41	0.59	0.40	1.00009	0.40
2001	2.08	5,319	13,325	0.40	0.60	0.40	1.00008	0.40
2002	1.88	6,745	14,683	0.46	0.54	0.36	1.00007	0.36
2003	1.91	7,889	16,818	0.47	0.53	0.36	1.00006	0.36
2004	2.05	8,025	18,329	0.44	0.56	0.39	1.00005	0.39
2005	2.04	8,596	19,668	0.44	0.56	0.38	1.00005	0.38
2006	2.10	9,149	20,845	0.44	0.56	0.38	1.00005	0.38
2007	**1.80**	11,578	21,752	0.53	**0.47**	0.30	1.00005	0.30
2008	2.08	10,583	22,750	0.47	0.53	0.35	1.00004	0.35
2009	2.01	12,421	24,289	0.51	0.49	0.33	1.00004	0.33
2010	1.92	12,898	26,222	0.49	0.51	0.32	1.00004	0.32
2011	2.06	11,556	28,367	0.41	0.59	0.37	1.00004	0.37
2012	**2.26**	10,527	31,044	0.34	**0.66**	0.42	1.00003	0.42
2013	2.11	13,526	32,297	0.42	0.58	0.37	1.00003	0.37
2014	2.02	15,520	35,430	0.44	0.56	0.35	1.00003	0.35
2015	2.09	14,874	34,847	0.43	0.57	0.36	1.00003	0.36
Total	2.04	1,63,972	3,72,260	Average	0.56	0.37		

Note: Cells in the highlighted font shows the highest and lowest values; SA= Literature of Single author; TL= Total Literature

The DC values vary from 0.47 to 0.66 with an average of 0.56 which indicates that there exists a moderate level of degree of collaboration among authors in the field of public health. The value of CC does not represent high collaboration among authors of public health from 2000 to 2015. The highest CC has been observed in 2012 (0.42) and the lowest one is 0.30 in 2007. The CC for public health authors lies between 0.30 and 0.42 with an average of 0.37 which means there is no significant magnitude of collaboration among the authors during the study period. As there is the existence of frequencies in the case of single authorship under the present study, the values of RCC is equivalent to the values of CC.

5.2.3 Authors' productivity

Yoshikane et al (2009) revealed the diachronic correlation of properties to measure the author's productivity by devising formulas that had been slightly modified by Mamdapur et al (2014)in their work.

$$AAPP = \frac{TA}{TP} \qquad \textbf{[Eq. 9]}$$

$$PPA = \frac{TP}{TA} \qquad \textbf{[Eq. 10]}$$

Where, AAPP = Average Author Per Paper
PPA = Productivity Per Author
TA = Total number of Authors
TP = Total number of Publication

Table 9: AAPP and PPA in public health literature

Year	Total Authors	Total Publications	AAPP	PPA
2000	25,035	11594	2.16	0.46
2001	27,678	13325	2.08	0.48
2002	27,627	14683	1.88	0.53
2003	32,165	16818	1.91	0.52
2004	37,557	18329	2.05	0.49
2005	40,040	19668	2.04	0.49
2006	43,716	20845	2.10	0.48
2007	39,151	21752	**1.80**	**0.56**
2008	47,342	22750	2.08	0.48
2009	48,882	24289	2.01	0.50
2010	50,277	26222	1.92	0.52
2011	58,421	28367	2.06	0.49
2012	70,221	31044	**2.26**	**0.44**
2013	68,271	32297	2.11	0.47
2014	71,679	35430	2.02	0.49
2015	72,749	34847	2.09	0.48
Total	7,60,811	372260	**2.04**	0.49

Note: Cells in the highlighted font shows the highest and lowest values.

The average author per paper is the value equivalent to CI (Collaborative Index) (Table 8). It is noted in Table 9 that the average author per publication is 2.04 means there are more than two authors per paper during the period 2000-2015. The average productivity per author (PPA) is 0.49 which means every author produces less than half of a publication each year during the study period. The average production rate per author ranges between 0.44 and 0.56. In 2007 authors had high production rate (0.56) whilst the year 2012 was the lowest productive year from PPA point of view (0.44).

To analyze the relationship between collaborative authors (X) and their productivity (Y) the null hypothesis was tested using a parametric test (has been excluding anonymous and single authors and their productions). Table 10 shows that the squared R-value reflects the similarity in the distribution overall which is equivalent of Pearson's r for two sets of values. The value of R Square indicates that productivity is related to collaborative authors. In this case, the p-value is significantly less than 0.01 meaning the regression coefficient is statistically significant. In summary, author productivity is influenced by the collaboration of authors and there is a strong positive correlation between authors' collaboration and authors' productivity.

Table 10: Regression coefficient

Regression Statistics			***Coefficients***	***Standard Error***	***t Stat***	***P-value****
Multiple R	0.996197823	Intercept	-582.1885122	322.3823316	-1.805894602	0.092475209
R Square	0.992410103	Author collaboration	0.348315067	0.008141053	42.78501359	**3.04948E-16**
Adjusted R Square	0.991867968					
Standard Error	432.8166858					
Observations	16					

Note: *significant at p<0.01

6. Further study and conclusion

The present study was designed to assess the growth pattern, author collaboration and productivity of public health literature in the period 2000-2015. During the study period in question (2000-2015), using Scopus database, 3,72,260 documents were identified, 2014 is the most productive year in terms of the highest number of the publication produced and the year 2000 was the least productive. The future growths of literature on public health were predicted using time series analysis. The number of publications in the literature on public health was 11,594 in 2000, which is predicted to grow to 61,902 in 2032. CC values on public health literature lie between 0.30 and 0.42 with an average of 0.37 which means there is no significant magnitude of collaboration among the authors during the study period. The PPA of public health literature ranges between 0.44 and 0.56 with an average of 0.49 which means that public health authors produce less than half a publication each year during the study period in question.

There is a trade-off relation between RGR and Dt(a) values. If RGR value increases Dt(a) value goes down. That means if the RGR of particular literature for a specific period increases, it will take less time to double. Otherwise, it will take more time to be doubled. The value of CC indicates whether the collaboration between authors on a particular subject in a specified period is high or low. CC as a number lies between 0 and 1. After calculating Collaborative Coefficient it can be said that if the value of CC is closer to '1' it means there is the high collaboration of authors and when it is closer to '0' indicating a weaker collaboration of authors. The value of CC value can't be interpreted the exact level of collaboration i.e. a particular CC value indicate which level of author collaboration is not specified previously. The interpretation of CC value to the exact level of collaboration could be an interesting area of future research. In measuring collaboration the inequality of values for CI, DC, and CC has been observed. It is noted that the value of CI is always greater than DC, which is further greater than the CI. So we can say easily for the same set of values that the equation would be CI > DC > CC. The assessment of public health literature with the help of these indicators could be very useful to researchers, scientists, library and information professionals, policymakers, and government agency relating to the concerned fields.

References

Ajiferuke, I. Burell, Q. and Tague, J. (1988). Collaborative coefficient: a single measure of the degree of collaboration in research. *Scientometrics*, Vol. 14 No. 5-6, pp. 421-433.

Arya, C. and Sharma, S. (2012). Authorship trends and collaborative research in veterinary sciences: a bibliometric study. *Chinese librarianship: an international electronic journal*, Vol.37, Retrieved from www.iclc.us/cliej/cl34AS.pdf

Balasubramani, C. (2015). *Research out put on genetic engineering: a scientometric study. (*Unpublished doctoral thesis), Bharathiar University, Coimbatore, India.

Brookes, B. C. (1990). Biblio-, Sciento-, Infor-metrics??? What are we talking about? In: EGGHE, L. & ROUSSEAU, R. (Eds), Second International Conference on Bibliometrics, Scientometrics and Informetrics. Informetrics 89/90 (pp. 31–43). Amsterdam: Elsevier.

Chuang K-Y., et al. (2011). Bibliometric analysis of public health research in Africa: The overall trend and regional comparisons. *South African Journal of Science*, 107(5/6). DOI:10.4102/sajs. v107i5/6.309

Clarke, A. et al. (2007). A bibliometric overview of public health research in Europe. *European journal of public health*, 17 (1).DOI:10.1093/eurpub/ckm063

Ding, Y., Foo, S., & Chowdhury, G. (1999). A bibliometric analysis of collaboration in the field of information retrieval. *The international information & library review*, 30, 367-376.

Donner, P., Chi, Pei-Shan, & Aman, V. (2014). Bibliometric study for German National Academy of Sciences Leopoldina in disciplines public health and epidemiology. Berlin, Germany: InstitutfürForschungsinformation und Qualitätssicherung (iFQ).

Egghe, L. (1991). Theory of collaboration and collaborative measures. *Information processing & management*, 21 (2/3), 177-202.

Elango, B., & Rajendran, P. (2012). Authorship trends and collaboration pattern in the marine sciences literature : a scientometric study. *International Journal of Information Dissemination and Technology,* 2(3), 166-169.

Gajbe, S. S., & Sonawane, S. S. (2015). Authorship pattern and degree of collaboration in the leprosy research: a scientometric study. *Knowledge librarian*, 2(6).

Garfield, E. (2009). From the science of science to Scientometrics visualizing the history of science with HistCite software. *Journal of Informetrics, III* (3), 173–179.

Glänzel, W. (n.d.). *A concise introduction to bibliometrics & its history*. Retrieved January 25, 2017, from Ecoom: https://www.ecoom.be/en/research/bibliometrics

Heidari, M., & Safavi, Z. (2013). The survey of collaborative coefficient of article authors in "Iranian Journal of Pathology" since 2006 to 2012. *Iranian Journal of Pathology*, 8(3), 165-170.

Hemala, K., & Kavitha, E.S. (2016). A scientometric analysis of autism productivity (2011-2015): study based on web of knowledge database. *International journal of information sources and services*, 3(6), 44-54.

Hood, W. W., & Wilson, C. S. (2001). The literature of bibliometrics, scientometrics,and informetrics. *Scientometrics, 52* (2), 291–314.

Hunt, R. (1982). Plant growth curves. London:Edward Arnold.

Hunt, R. (1990). Relative growth rates. In R. Hunt, *Basic Growth Analysis* (pp. 25-34). Springer.

Jimenez-Fanjul, N., Maz-Machado, A., & Bracho-Lopez, R. (2013). Bibliometric analysis of the mathematics education journals in the SSCI. *International journal of research in social science*, 2(3), 26-32.

Kalita, A., Shinde, S., & Patel, V. (2015). Public health research in India in the new millennium: a bibliometric analysis. Global health action, 8. DOI: http://dx.doi.org/10.3402/gha.v8.27576

Keshava (2004). *Scientometric analysis of social science research in India* (Unpublished doctoral thesis). Karnatak University, Karnataka, India.

Kumar, R.S. & Kaliyaperumal, K. (2015). A scientometric analysis of mobile technology publications. *Scientometrics*, 105:921-939.

Kumar, S., & Naqvi, S.H. (2014). Collaboration pattern in the field of natural sciences at Jamia Millia Islamia, New Delhi during 1971-2007. 9th convention PLANNER. Assam: INFLIBNET.

Lawani, S.M. (1980). Quality, collaboration, and citations in cancer research: a bibliometric study (Unpublished doctoral thesis). Florida: Florida State University, USA.

Macias-Chapula, C.A. et al. (2008). A bibliometric analysis of public health research work in Mexico. In Kretschmer, H. & Havemann, F. (Eds), Fourth International Conference on Webometrics, Informetrics and Scientometrics & Ninth COLLNET Meeting. Berlin: Gesellschaft¨ur Wissenschaftsforschung.

Mahendran, M. (2014). *Scientometric analysis of epidemiology research output: a global perspective* (Unpublished doctoral thesis). Annamalai University, Annamalai Nagar, India.

Mamdapur, G. M. N. (2014). Bibliometric Portrait of SRELS Journal of Information Management for the Period 2004-2013. *Library Philosophy and Practice (e-journal). Retrieved from* http://digitalcommons.unl.edu/libphilprac/1166

Manickaraj, J. (2015). *Scientometric analysis of literature on wireless communication (1998-2012) using scopus database* (Unpublished doctoral thesis). Bharathiar University, Coimbatore, India.

Manimegalal, M. (2014). *Growth and collaboration trends in the field of fashion technology: a scientometric analysis* (Unpublished doctoral thesis). Annamalai University, Annamalai Nagar, India.

Meena, M. (2015). *A Scientometric* analysis *of the trends in the malaria research* (Unpublished doctoral thesis). Annamalai University, Annamalai Nagar, India.

Meera, S. (2007). *Indicators in the field of chemical sciences with special reference to India: a scientometric study* (Unpublished doctoral thesis). Karnatak University, Karnataka, India.

Merigo, J.M. & Nunez, A. (2016). Influential journals in health research: a bibliometric study. *Globalization and health*, 12(46). DOI 10.1186/s12992-016-0186-4.

Murphy, C.B. (2019). Compound Annual Growth Rate-CAGR. *Investopedia.* Retrieved from https://www.investopedia.com/terms/c/cagr.asp

Murugiah, P. (2015). *Research productivity of Human DNA literature: a scientometric analysis* (Unpublished doctoral thesis). Alagappa University, Tamil Nadu, India.

Nalimov, V., & Mul'chenko, Z. M. (1969). Naukometriya. Izuchenie Nauki kak Informatsionnogo Protsessa (Scientometrics. the study of science as an information process). Moscow: Nauka.

Narayanan A L, B. (2015). *Scientometric analysis of literature on Mems: a study based on Scopus database* (Unpublished doctoral thesis). *Anna university*, Chennai, India.

Packiyaraj, M. (2014). *Scientometric analysis of literature on textile technology based on bibliographic databases* (Unpublished doctoral thesis). Bharathiar University, Coimbatore, India.

Parker, R. (2002). Planning analysis: calculating growth rates. Retrieved August 26, 2019, from http://pages.uoregon.edu/rgp/PPPM613/class8a.htm

Pillai, K.G.S. (2007) Authorship patterns in physics literature: an informetric study on citations in doctoral theses of the Indian Institute of Science. *Annals of library and information studies*, 54, 90-94.

Prabakar, S. (2014). *Scientometric analysis on literature output on nano thin films from secondary database Web of Science* (Unpublished doctoral thesis). Manonmaniam Sundaranar University, Tirunelveli, India.

Pritchard, A. (1969). Statistical bibliography or bibliometrics? *Journal of Documentation, 24*.

Rakhi, VS. (2014). Collaboration trend in Indian business-management research: a bibliometric perspective. Library philosophy and practice. Retrieved August 26, 2017, from http://digitalcommons.unl.edu/libphilprac/1139

Ramesh, P. (2015). *Research out put on brain tumor: a scientometric study* (Unpublished doctoral thesis). Bharathiar University, Coimbatore, India.

Rao, M.K. (2005). *Evolution of materials science and engineering: a scientometric study* (Unpublished doctoral thesis). University of Madras, Chennai, India.

Ravanan, C. (2012). *Scientometric analysis of agriculture literature: a global perspective* (Unpublished doctoral thesis). ManonmaniamSundaranar University, Tirunelveli, India.

Sachithanantham, S. (2017). *Current trend on Rabies research publication: a scientometric study* (Unpublished doctoral thesis). Alagappa University, Tamil Nadu, India.

Sahoo, B. B. (2006). *A scientometric study of literature in software studies in India with comparison to world literature* (Unpublished doctoral thesis). University of Pune, Pune, India.

Sandha, D.P. (2001). *Metallurgy and material sciences research in India (1985-1994)* (Unpublished doctoral thesis). Utkal University, Orissa, India.

Savanur, K., & Srikanth, R. (2010). Modified collaborative coefficient: a new measure for quantifying the degree of research collaboration. *Scientometrics,* 84.

Sedam, M. V. (2013). *Scientometricportraiots of selected nobel laureates in the field of Physics* (Unpublished doctoral thesis). Karnatak University, Karnataka, India.

Senapati, S.K. (2009). *Bibliographic control of periodical literature on building materials published in India: a scientometric study* (Unpublished doctoral thesis). Punjabi University, Patiala, India.

Sivakami, N. (2015). *A scientometric analysis of swine influenza research output from 1990 to 2013* (Unpublished doctoral thesis). Alagappa University, Tamil Nadu, India.

Subramanyam, K. (1983). Bibliometric studies of research collaboration: a review. *Journal of information science*, 6, 33-38.

Tejashwini, B. (2016). *Social Network analysis of author collaboration in Biotechnology: a study* (Unpublished doctoral thesis). U*niversity of Mysore*, Mysore, India.

Thilakar, S., & Ponnudurai, R. (2013). Contributions to crop science research: measuring authorship pattern. *International Journal of Humanities and Social Science Invention*, 2(12), 35-40.

Todeschini, R., & Baccini, A. (2016). Handbook of bibliometric indicators: Quantitative tools for studying and evaluating research. Germany: Wiley-VCH.

Udofia, U. I. (2002). Bibliometric studies on African Trypanosomiasis research literature: a review of authorship pattern. *Annals of library and information studies*, 49 (2), 45-49.

Venkatesan, M.N. (2015). *Nuclear power generation research output: a scientometric analysis* (Unpublished doctoral thesis). Alagappa University, Tamil Nadu, India.

Yazit, N. & Zainab, A.N. (2007). Publication productivity of Malaysian authors and institutions in LIS. *Malaysian journal of library & information science*, 12(2), 35-55.

Yoshikane, F. et al. (2009). An analysis of the connection between researchers' productivity and their co-authors' past contributions, including the importance in collaboration networks. Scientometrics, 79(2), 435-449.