



Visualizing the knowledge outburst in global research on COVID-19

Jiban K. Pal¹

Received: 9 September 2020 / Accepted: 8 February 2021 / Published online: 6 March 2021
© Akadémiai Kiadó, Budapest, Hungary 2021

Abstract

The scholarly output of the new coronavirus research has been proliferating. During five months, an amount of 14,588 scientific publications about nCoV-2 and COVID-19 has been generated intensively (as indexed in Scopus on 31 May 2020). Such a knowledge outburst has created ample interest in understanding the research landscape of this newly configured area. This paper demonstrates on scientometric dimensions of the novel coronavirus (2019-nCov) research using quantifiable characteristics of the publication dataset. Findings reveal that the rate of publication growth (1600%) is very significant to a synergic response of the researchers to combat with the most extended sequence of an RNA virus. Indeed their response has geared up to an average of 100 articles per day. Many scholarly publishers have disclosed their preprint servers to make the publications available immediately, even by enabling Open Access. The scientific contents have published in more than 500 journals from 240 academic publishers. While the top-ten publishers occupied almost 70% of the articles, then about 25% of the studies were sponsored by 300 funding agencies. Among the notable journals Lancet, Nature, BMJ, JAMA, JMV, and NEJM are prominent. Findings also reveal that majority of the contributions have occurred in Medical Science, focusing on virology, immunology, epidemiology, pharmacology, public health, critical care, and emergency medicine. However, the closely associated terms are virus transmission, infection control, asymptomatic, quarantine, pneumonia, human, disease severity, clinical trials, viral pathogenesis, pandemic, risk, and mortality. The study suggests that academic hubs are located mostly in the USA, China, Italy, and the UK. Among the productive institutions; Huazhong Univ (China), Tongji Med. College (China), Harvard Med. School (USA), Univ of Milan (Italy), INSERM (France), UCL (UK) are outstanding. The G7 countries together produced 50% of the global research output on nCov-2. It also noted an encouraging trend of collaborative research across many countries and disciplines, where the values of CI (6.46), DC (0.79), and CC (0.59) are very significant. It examines the geographical diversity of the collaborating authors, thereby visualized their linkages via co-authorship occurrences. Finally, it analyzed the publications' impact to showcase the most influential contributions of the new coronavirus research.

Keywords Scientometrics · Scientific visualization · Research evaluation · Knowledge mapping · Quantitative analysis · Novel Coronavirus · COVID-19 · 2019-nCoV · SARS-CoV-2

Introduction

A pandemic situation around the Novel Coronavirus is unprecedentedly affected our life and livelihood. The virus first struck in the Wuhan city of China by the end of December 2019. Then it has spread very rapidly and widely (through travelers) across many regions, particularly in Europe and America. Now, almost all the countries are grappling with this deadly Corona Virus Disease (COVID-19). There have been 5,891,182 confirmed cases of infection, including 365,966 fatalities, occurred globally as reported by WHO on 31 May 2020 (<https://covid19.who.int>). Most of the countries are experiencing an extended lockdown. The global economy has fallen into a recession. The job market is gradually losing its temper worldwide. So this eruption has become a buzzword in every sphere of civilized society.

The agenda has touched the lives of every human being on Earth, bringing out the best and worst of human emotions, among others. In this challenging time, scientists around the world have been expressing their concern and workmanship through academic cohesiveness. The phenomenal increase of scholarly literature on coronavirus disease has been observed. The Google Search yielded about 458×10^7 results of the on 31 May 2020). Scholar-google also returned a total of 147,000 hits on the topic. A retrieval of this keyword in the re3data (a global registry) has resulted as many as 34 research data repositories. However, the WHO COVID-19 Database (www.who.int) has recorded more than 22,000 entries of current literature from multilingual sources (with a few overlaps). Approximately 14,588 scholarly articles are indexed in Scopus (www.scopus.com) during a few months of 2020.

Such an overwhelming growth of scientific literature has created sufficient interest in visualizing the knowledge base of this promising issue. A sudden outburst in publishing activity surrounding the topic of COVID-19 has occurred. Indeed a collective agenda of global research posed by the researchers (with utmost attention) prompted us to address their contributions scientometrically. Otherwise, no such work has been found in terms of its' scope and objectives. So, it was felt worthy of looking at the publication scenario of this specialty of research in different dimensions.

Earlier efforts

A few studies to map the research on novel coronavirus disease (COVID-19) have been evident in the literature. An earlier attempt to measure the new coronavirus research was made by Lou et al. (2020) using 183 publications indexed in the PubMed database until 29 February. They analyzed and reviewed the papers about COVID-19 on different aspects of bibliometric variables. Nasab and Rahim (2020) also examined the scientific outputs of COVID-19 obtained from multiple sources (viz. Web of Science, Scopus, and PubMed) as found on 10 March 2020. In their study, they considered a total of 227 publications to assess in terms of authors, keywords, and cited-references. Chahrour et al. (2020) evaluated the research activity on this killer demon correlating the GDP rates and population of 39 countries. They have extracted the publication data from PubMed and WHO-COVID-19 dashboard as of 18 March. Hossain (2020) studied on the publications of COVID-19 as reflected in the Web of Science (core collection) on 1st April. The study identified a dataset of 422 publications to map the areas of knowledge using bibliometric tools. In a study, Dehghanbanadaki et al. (2020)

counted the citations of 923 research publications on COVID-19 found in Scopus by early April.

Besides that, some other studies have been conducted to visualize the results from different perspectives. Zhang et al (2020) attempted to quantify and characterize the responses of academia to public health emergencies since 2000. They have identified the five outbreaks of four viruses (SARS, H1N1, Zika, Ebola), thereby compared the activity patterns with the Covid-19 pandemic. Their study analyzed the publication dynamics in terms of geographic region, institutional sector, and funding agency using comparative bibliometric approaches. Teixeira da Silva et al (2020) noticed that there is a huge gap between the Covid-19 publication records covered in Web of Science and Scopus. In the first six months (1 January to 30 June 2020) about 12,331 and 21,602 publications were retrieved from Web of Science and Scopus, respectively. In their letter, they remarked on some other aspects with a few limitations of the dataset.

Belli et al (2020) also analyzed the scientific publications of new coronavirus in terms of international collaboration and open access those have been characterizing the open-science. They mapped and compared the publication dataset covered in Web of Science (core collection) from 2001 to 2020. Deng et al (2020) examined more than 15,000 publications (in English and Chinese) related to ‘human coronaviruses’ as indexed in Scopus for a period from 1960 until 15 February 2020. Their study was intended to provide information about the pathogen as a whole (but not limited to the new coronavirus) keeping in view to map the previous research in this field. Haghani and Bliemer (2020) pursued a scientometric comparison across the literature on SARS, MARS, and nCoV-2. Their study focused on the co-occurrence of keywords, citation relations of journals, bibliographic coupling along with the international collaboration scenario. Summarily they found that the rate of scholarly output (prompted by 2019-nCoV) is unique in the history of scientific publication.

More recently, Yu et al (2020) studied with a small dataset comprising 3626 papers about COVID-19 published between December 2019 and May 2020. They have conducted a search (applying “COVID-19 or Novel Coronavirus or SARS-CoV-2 or 2019-nCoV” keywords in the title) to retrieve the publication records from Web of Science on 20 May 2020. Retrieved data were analyzed in terms of 4Cs (Co-occurrence, Co-authorship, Co-citation, and bibliographic Coupling) using weight attributes of the VOSviewer software.

The present study has considered a distinct dataset using a more powerful search string to focus on the knowledge outburst of scientific publications about nCoV-2 and COVID-19. It has analyzed an amount of 14,588 papers generated during the first five months of the global health emergency for COVID-19. Thus it demonstrates the synergic response of global researchers to visualize the research landscape of this newly configured area.

So this study is quite different from earlier efforts in terms of its scope & objectives and certainly unique in using a different set of indicators for scientometric analyzes. Indeed, no such intrinsic study has been made earlier to depict the synergic growth and orientation of scholarly behavior, thereby many other aspects of the new coronavirus research—*hence this sincere effort*. It will be helpful to propose future strategies to elevate the research in this emerging area for the prevention and control of COVID-19.

Aims and objective

This paper aims to visualize the incidental break-out of scientific information concerning the new coronavirus disease (COVID-19). It also aims to showcase the scientometric dimensions of novel coronavirus (2019-nCov) research. Therefore, this study intends to map the scholarly publications of this promising issue in quantitative terms. The real spirit of this work is implicit in a curious response of the global researchers during a public health emergency.

Objectively, this study has been conducted to realize the orientation of COVID-19 research for providing informational support to the researchers and academic administrators thereof. It is therefore observed the publication scenario in order to empower the research governance to be warranted in this scientific specialty. The academic directors can be able to identify their peers, policymakers could find the strategically relevant areas of research, and funding bodies could make a more informed decision for their investments. Thus, it may be considered as an instrument for capacity building and could be utilized to enrich the understanding of stakeholders for enabling better utilization of resources and facilities toward an immense scientific interest.

Otherwise, the study is conducted with the following objectives;

- a) To understand the quantum of knowledge through scholarly publication growth in COVID-19.
- b) To observe the orientation of research and to visualize the corpus of research focus.
- c) To identify the prolific contributors and profound institutions of this research.
- d) To determine the extent of collaborative research across the countries.
- e) To map the collaboration network through co-authorship analysis.
- f) To reveal the dynamics of publishing and find core-journals.
- g) To evaluate the influence of publications as a whole.

Materials and method

In order to pursue the above objectives, required data for this study have been collected from Scopus. The Scopus® is a widely used bibliographical database having a powerful search interface with the broadest coverage and timely updates. So the publication records of scholarly articles relating to COVID-19 were first retrieved from the source database, subsequently extracted (in CSV format), and downloaded with all bibliographic elements.

Noteworthy is the fact, among many other databases; Scopus is the most comprehensive source of scientific publications from diverse areas of research. Indeed the coverage of Scopus always supersedes major databases. Teixeira da Silva et al (2020) found that there is a huge gap between the publication records of nCoV-2 covered in WoS and Scopus. Thus it alone qualifies to hold a gigantic proportion of global research (representative sample) and can be a reliable source of current publication data. Again, the use of parallel two or more databases often generates too many false-positives (item overlaps) within the dataset. Therefore, it is well-recommended by evaluative scientometricians to consider the Scopus as the single greatest source of publication data.

So the global researches about COVID-19 published from January to May 2020 were retrieved through advanced search interface of the Scopus database. Boolean operators

are used for combining different search results of a few possible terms with the following search expression: [TITLE-ABS-KEY (“COVID-19” OR “COVID19” OR “COVID-2019” OR “2019-nCov” OR “2019 Novel Coronavirus” OR “Coronavirus Disease 2019” OR {SARS-CoV-2} OR “Coronavirus”) AND PUBYEAR = 2020], applied to the Title, Abstract, and Keyword (TITLE-ABS-KEY) fields. The syntactical use of parentheses and double-quotes in the search string made this formulation more strategic for a precise result with optimum potential. The searching reliably covered all types of articles published during a period of five months without limiting any language.

The final search yielded 14,588 publication records (until 31 May 2020)—an interesting sample size within a limited period (Fig. 1). The retrieved records were then converted to Microsoft Excel and organized them methodically. MS-Excel provided excellent facilities for analyzing the dataset.

The study is, therefore, designed empirically to analyze the scholarly publications of ‘new coronavirus’ research produced during the first five months of the global health emergency due to Covid-19. The analyses have been made on the basis of quantifiable characteristics of the publication dataset (as reflected in Scopus) using scientometric methods and techniques.

Ultimately, required computations are carried out to make observations and drawing inferences. Descriptive statistics are used to present the quantum of information meaningfully. The growth of publications and decadal changes over a passage of time has been plotted. Scopus subject categories are used to obtain the major aspects (subject-clusters) of this research. For visualizing the prominent topics of research, author-generated keywords are analyzed through VOSviewer (ver.1.6.13) software. The relative potency of different countries in the new coronavirus research is also analyzed. Rankings of productive authors, institutions, and countries are made by the authorship occurrences (counted straightway). Collaboration strength has been determined by the Degree of Collaboration (Subramanyam, 1983), Collaborative Index (Lawani, 1980),

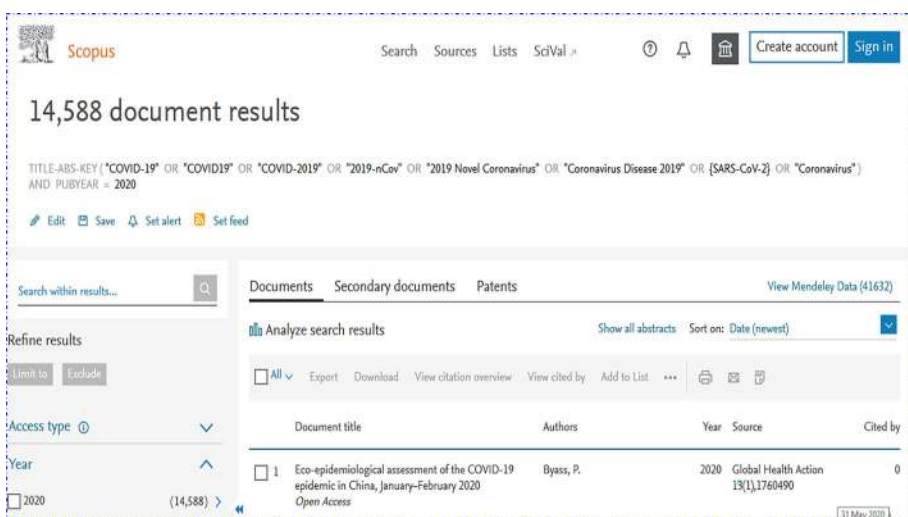


Fig. 1 Search results returned from Scopus (on 31 May 2020)

and Collaborative Coefficient (Ajiferuke et al., 1988). The collaboration network of the countries is also visualized through co-authorship analysis.

Core journals are identified by tracking the preferred channels of research communications. In ranking core journals, a new measure combining both ‘Impact Score of a journal’ (viz. Impact Factor or CiteScore) and ‘Frequency of relevant publications’ has been employed. Such a metrics can provide more rational and trusted view to assess the value of a source-title through weighted score, which may be utilized to identify potential journals with finer tunings. While measured the influence of the publications, a trade-off between the quality and quantity has occurred through citations received from Scopus as of 31 May 2020.

Limitations of the study

Scientometric evaluations cannot be a single study, rather “*a succession of studies that critique and improve upon each other*” (Feller et al., 1996). This study considered only scholarly articles relating to the new coronavirus (nCoV-2) research, as reflected in Scopus. So the less scholarly materials (like conference papers, technical reports, newspaper items, debate notes, interviews, and patents are omitted from the purview of this study.

However, the sheer volume of publications indexed in Scopus might offer more comprehensive coverage of global research output; but different databases have their lacuna (bias) in terms of choice-of-sources and update entries. This study provides a snapshot of the prototype dataset using Scopus only. Therefore, the restriction imposed in database use (Scopus) and time-span covered (five-months). So the variation in time-frame, as well as scope and coverage of the dataset, might lead to a difference in counting—hence, may be considered as a macro-level assessment.

Indeed this study is based on ‘quantitative’ measures using the ‘straight’ counting technique (rather than employing fractional scores) and free from qualitative evaluations. Even though, all publications do not contribute equally to the scientific knowledge and do not produce equal impacts on further research; but, equal values of all publications have been taken into account. Thus it imposed upon general limitations of scientometric studies.

Any conclusion is, therefore, subject to the above limitations and to be interpreted with caution. A more in-depth study may be needed to ascertain the results.

Analysis and findings

The empirical dataset revealed a lot of exciting results. It analyzes 14,588 publication records about the new coronavirus research (appeared from January to May in 2020) with a factual presentation of relevant data.

Growth of publications

The first impression of coronavirus research was found in 1951. Gradually it had created attention to the researchers when this viral-infection was transmitted to humans in 1965. Further growth of research has been noticed after the first strike of its infection caused Severe Acute Respiratory Syndrom (SARS) that emerged in China in 2002.

Table 1 Growth of publications in coronavirus research, 1951–2020

Year (Pubs)	Year (Pubs)	Year (Pubs)	Year (Pubs)	Year (Pubs)	Year (Pubs)
1951 (4)	1971 (18)	1981 (170)	1991 (162)	2001 (230)	2011 (517)
1957 (1)	1972 (33)	1982 (127)	1992 (155)	2002 (149)	2012 (577)
1962 (3)	1973 (31)	1983 (126)	1993 (167)	2003 (1009)	2013 (810)
1964 (6)	1974 (61)	1984 (161)	1994 (229)	2004 (1097)	2014 (939)
1965 (16)	1975 (80)	1985 (138)	1995 (253)	2005 (941)	2015 (1004)
1966 (18)	1976 (73)	1986 (136)	1996 (161)	2006 (784)	2016 (918)
1967 (18)	1977 (80)	1987 (228)	1997 (184)	2007 (640)	2017 (830)
1968 (10)	1978 (113)	1988 (139)	1998 (229)	2008 (628)	2018 (812)
1969 (17)	1979 (86)	1989 (158)	1999 (160)	2009 (576)	2019 (850)
1970 (24)	1980 (122)	1990 (249)	2000 (157)	2010 (563)	2020 (14,588)

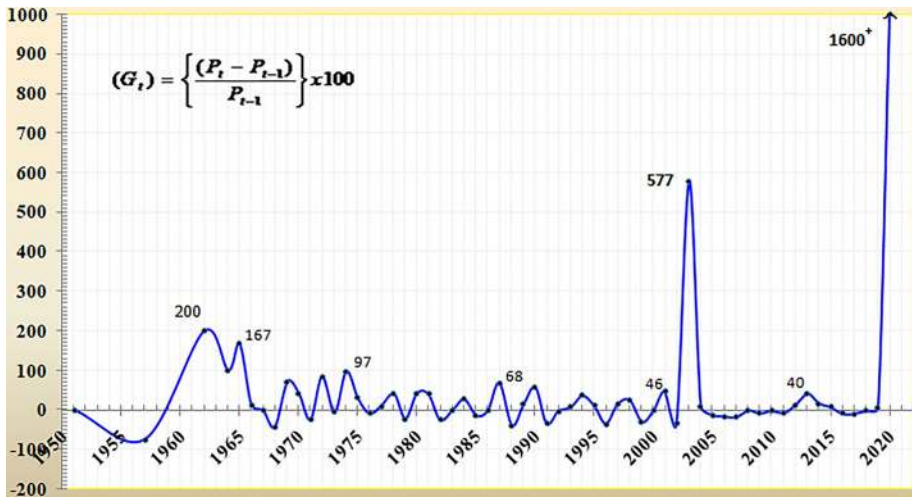


Fig. 2 Publication growth curve (rate in percent)

That epidemic spread across twenty-eight countries, with more than 8000 infected people, including 774 deaths within a year. However, the outbreak prompted researchers to consider it as a prominent issue, since 2003. An increased volume of publications with the big family of coronaviruses is evident. However, the new coronavirus (so-called SARS-CoV-2) that causes COVID-19 has created special attention to the researchers worldwide. So an outburst of knowledge has occurred in the year 2020. Such a synergic (furious) response of the global researchers to fight against the pandemic situation is unprecedented.

Table 1 (Fig. 2) depicts the growth of publications relating to coronavirus research over the years. Publication growth refers to the scientific progress that reasonably correlates the allocation of resources, funds, and individual grants for research and development.

The growth rate implies the percentage of change in research output over some time. Here, growth rate (G_t) is defined as the rate of increase (or decrease) of publication output (P) in the year (t) compared to the previous year ($t - 1$) and calculated by percent (%).

$$(G_t) = \left\{ \frac{(P_t - P_{t-1})}{P_{t-1}} \right\} \times 100$$

The growth curve follows an irregular pattern with its first spike (577%) in 2003 when the SARS-CoV-1 epidemic contaminated over 28 countries. The next surge of the publication growth (1600+ %) is significantly found at the time of the global pandemic due to SARS-CoV-2 (COVID-19). It is undoubtedly a knowledge outburst that occurred by producing about 14,588 scholarly publications within a few months in 2020.

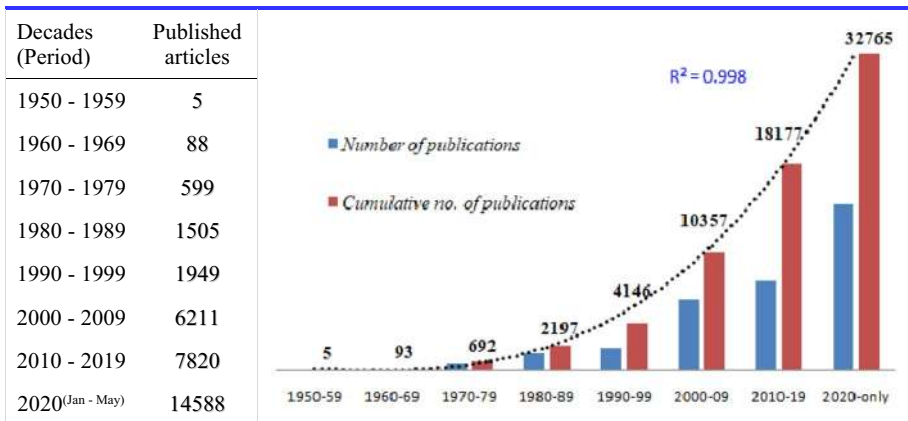
It also presents the publication growth of this specialty of research in the last seven decades. The growth trend is presented through a regression curve plotted with the cumulative distribution of publications. Table 2 shows an increasing order of growth trend following polynomial type regression ($R^2=0.998$). The publication-size for the year 2020 (Jan-May, 14,588) is certainly indicative. As such, the volume of publication is representing more than a decade. Such an exponential growth-trend has never been found in the history of research. Thus it bears a special significance to draw the insights on the knowledge-base of new coronavirus (nCov-2) research.

Orientation of research

Efforts have been made to realize how the research communications are oriented in different forms, languages, publishers and mechanisms of access imposed therein. However, these research endeavors (scientific contents) were sponsored by numerous funding-bodies, are also examined. A total of 14,588 publication records about the new coronavirus found in 2020 have been consulted.

Findings reveal that about 80% of the total publications are enabled with the ‘Gold Open Access’ (11,584 out of 14,588). It implies that the researchers are used to publish their papers in online open access journals. In contrast, 20% of articles have either ‘Green OA’ or entirely restricted access with different paywalls. Belli et al (2020) in their study noticed that open access to coronavirus publications is increased from 59% (2001–2019) up to 91% (in 2020). However, the most common form of communication was a full-length research article (about 48%) followed by scientific letters (21%), editorial material (11%),

Table 2 Decadal growth of publications in coronavirus research



research notes (10%), review papers (8%), and short surveys (1%). The rest of the publication types constituted only 1% (viz. erratum, data paper).

The publications have appeared in 20 different languages, of which 13,612 (93%) are in English, then in Chinese (467), French (200), German (191), Spanish (190), and others. More importantly, the information contents are published in a total of 525 journals (serials) from more than 240 publishing houses. Among the renowned publishers, Elsevier occupies the central place (holding 3500 articles). It is followed by NLM-Medline, Wiley-Blackwell, Springer-Nature, Lancet-Group, BMJ-Group, WB ~ Saunders, Taylor & Francis, Nature Research, American Medical Association, CUP, Academic Press, BioMed Central, Sage publications are prominent. Notably, the top-ten publishing group together constituted almost 70% of the total scientific contents in this specialty of research.

It is also found that more than 300 funding-agencies sponsored about 25% (counted 3600) of the research efforts (scientific contents) through individual grants. Major funding agencies, those have extended their support via research grants or fellowships are: National Natural Science Foundation of China, Beijing (420); National Institutes of Health, USA (172); National Basic Research Program of China—Program 973 (93); National Institute of Allergy and Infectious Diseases, USA (74); Wellcome Trust, London, UK (51); National Institute for Health Research, UK (50); Fundamental Research Funds for the Central Universities (45); Pfizer (32); Canadian Institutes of Health Research, Ottawa, Canada (31); Bill & Melinda Gates Foundation, USA (28); National Science Foundation (24) are to name a few of many others.

Focused areas of research

For examining the significant areas of research, a descriptive analysis of the publications has been made by the subject categories adopted in Scopus. However, author-generated keywords have become operative for identifying the prominent issues nurtured by the researchers as a pandemic response. The VOSviewer tool has been used to visualize the active topics of research through a co-occurrence map of the frequent keywords.

Findings ascertain the broad subject-clusters like Medicine (60%), Biochemistry & Molecular Biology (8%), Immunology and Microbiology (7%), Social Sciences (4%), Nursing (3%), Pharmacology (3%), Neuroscience (2.5%), Environmental Science (2%), Health Professions (2%), Multidisciplinary and others. Again the vibrant subdomains of Medicine to combat the diseases from this longest sequence (about 30,000 letters) of RNA virus are; Immunology, Virology, Epidemiology, Public health, Critical Care, and Emergency Medicine. The most active topics are COVID-19, Coronavirus, Betacoronavirus, SARS-CoV-2, Coronavirus infection, Virus transmission, Human, Viral pneumonia, Infection control, Quarantine, Disease severity, Infection risk, Mortality, Clinical trials, Viral pathogenesis, Pandemic, etc. (as presented in Fig. 3).

The visualization map shows a few clusters of keywords represented by circles of different colours. A larger circle implies greater weight of the item, and their associated strength links all the circles. While closely located items are strongly associated, then thicker lines have stronger linkage. The strength of co-occurrence links with other keywords has been calculated with a minimum of 100 occurrences in the threshold. The greatest total links of top-fifty terms are selected through the VOSviewer tool. There are four clusters of varying colours with distinct circles of different items. Some of them are linked very closely viz.

University of California, USA (115); University of Washington, USA (110); Imperial College of London (105); and National University of Singapore (105) are prominent among others. For avoiding an elongated list, only the institutes having a minimum of 100 contributions are mentioned here.

Prolific contributors

This study reveals a few productive contributors to the new coronavirus research. For identifying prolific contributors, the number of authorships (for each contributor) has been quantified by straight counting. Elisabeth Mahase (Medical Journalist of the BMJ, London) is found as the most prolific author, who appeared in the byline of 64 publications about nCov-2; followed by Viroj Wiwanitkit (Hainan Medical University, China & DY Patil Vidyapeeth Deemed University, Pune, India—57 papers); Giuseppe Lippi (University of Verona, Italy—33 papers); Beuy Joob (Medical Academic Center, Bangkok, Thailand—30 papers); Alfonso J. Rodriguez-Morales (University of Pereira, Colombia—24 papers); Gareth Lacobucci (The BMJ, London—23 papers); Ziad A. Memish (Ministry of Health, Saudi Arabia—22 papers); Mohamad Goldust (University Hospital of Basel, Switzerland—21 papers); Abi Rimmer (The BMJ, London—21 papers); Torello M. Lotti (Sapienza University of Rome, Italy—20 papers) are to name a few of the others.

Notably, top-75 authors have produced a minimum of 10 publications each. They are a productive group of researchers for nCov-2. Within the top-hundred contributors, Chinese authors are prevalent. Among the Indian Scientists Raju Vaishya (Indraprastha Apollo Hospital, New Delhi – 14 papers), Anoop Misra (Fortis C-DOC Center for Diabetes, New Delhi – 12 papers), Rimesh Pal (PIMER, Chandigarh – 10 papers), Kuldeep Dhama (Indian Veterinary Research Institute, Bareilly), Nivedita Gupta (ICMR, New Delhi), Awadhesh Kumar Singh (G.D Hospital & Diabetes Institute, Kolkata), Ritesh Gupta (Fortis C-DOC Center for Diabetes, New Delhi), Abhishek Vaish (Indraprastha Apollo Hospital, New Delhi), and Debanjan Banerjee (NIMHANS, Bangaluru) are prominent.

Collaboration scenario

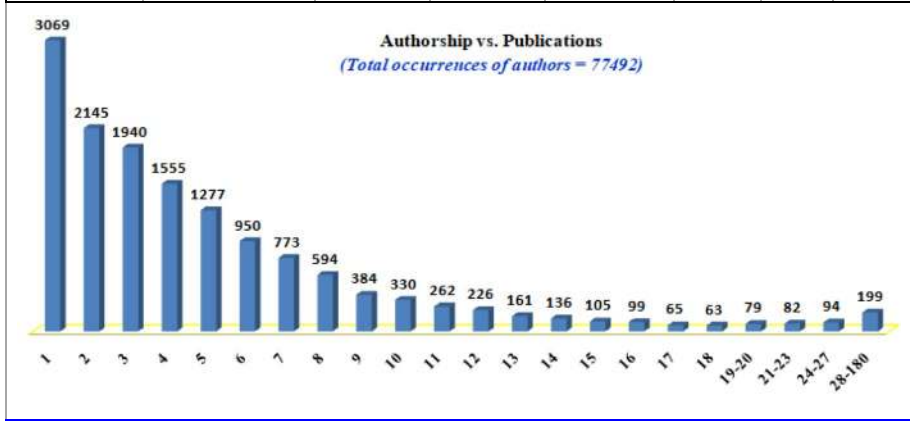
Collaboration is an effective way of sharing competencies and a considering factor of research productivity. Applied areas of research are prone to collaborative efforts. Quite often, the scholars are being forced to collaborate due to interdisciplinary agenda of research that requires knowledge across many fields (Pal, 2015). However, the positive impact on the citation of multi-authored publications has been promoting the growth of collaborations.

The collaboration scenario can be observed by occurrences of (co)-authorship in different degrees. Primarily it can be of two categories; namely Collaborative (N_m), and Non-collaborative (N_s). Most of the publications are Collaborative (11,519, 79%), and others were published by Single-authorship (3069, 21%). Further distribution of collaborative publications can be made in two distinct levels – viz Multi-authorship (having 2–9 authors), and Mega-authorship (10 or more authors). Table 3 depicts the distribution of authorships and the extent of collaborative research.

About 9618 (66% of total) papers are published under multi-authorship and mega-authorship publications are counted almost 13% (1901), where ten papers have more

Table 3 Authorship distribution and extent of collaborative research

Total publications	Single-authored publications (N _s)	Collaborative (N _m)		Authorship occurrences in N _m	CI	DC	CC
		Multi	Mega				
14588	3069	9618	1901	74423	6.46	0.79	0.59



than 100 authorships (each). When multi-authorship publications are significant, then the mega-authorship is evident up to 180 occurrences of authors. Mega-authorship publications together constituted a total of 32,389 (about 42% of the total) occurrences of authorships. A total of 77,492 occurrences of authorships are found in 14,588 publications considered for this study. Thus average authorship is computed (≈5) per paper.

To measure the extent of collaborative research, the Degree of Collaboration (DC) and Collaborative Index (CI) have been estimated. DC is an effective measure of collaborative efforts in quantitative terms (Subramanyam, 1983). It can be derived by the ratio of collaborative publications (N_m) and the total number of publications (N_m + N_s) at a given time.

$$DC = \frac{N_m}{N_m + N_s} = \frac{11519}{11519 + 3069} = 0.79$$

To determine the CI, the mean number of authors per collaborative publication is calculated (Lawani, 1980) using the formula given below. The CI is computed as 6.46 authors per collaborative publication (N_m).

$$CI = \frac{\sum_{j=1}^A j \cdot f_j}{N} = \frac{74423}{11519} = 6.46$$

While the DC (0.79) implies the prevalence of team-research (i.e., 79 papers out of 100), then the CI (6.46) indicates that the team-efforts are typically made by more than 6 (six) authors in the new coronavirus research. As a comprehensive measure and better alternative of DC and CI, the Collaborative Coefficient (CC) can be used to determine the collaboration strength (Ajiferuke et al., 1988). When the value of CC tends to '0' (as 1 - 1/j = 0), the publications are mostly non-collaborative. In reverse, the value when nearer to 1 (0.6) can be significant to a large number of collaborative efforts made by the scientists.

$$CC = 1 - \frac{\sum_{j=1}^k (1/j)F_j}{N}$$

where (F_j) is the number of j -authored publications within the total (N) publications and (k) denotes the highest number of authors for each publication. More simply, it can be expressed as;

$$CC = 1 - \frac{f_1 + (1/2)f_2 + (1/3)f_3 + \dots + (1/k)f_k}{N}$$

Putting the values of authorship and corresponding numbers of papers viz. solo-authorship (3069), two-authored (2145), three (1940), and so on from Table 3; the value of the CC (0.59) is obtained very significantly.

$$\begin{aligned} &= 1 - \frac{(1 \times 3039) + (1/2 \times 2145) + (1/3 \times 1940) + \dots + (1/180 \times 1)}{14588} \\ &= 1 - \frac{(3069) + (1072.5) + (646.66) + \dots + (0.0055)}{14588} \\ &= 1 - \frac{(5952.3887)}{14588} = (1 - 0.408033) = 0.59 \end{aligned}$$

However, the geographical diversity of collaborating authors is worth mentioning. The co-authorship network among the collaborating countries has been mapped based on link-strength. A total of 35 countries meet the threshold of a minimum of 50 collaborations each. Subsequently, the cumulated strength of co-authorship links for each of the countries is computed. Most of the collaborations have occurred between the USA, UK, Italy, China, and Australia. It means the profound collaborators are mostly located in these countries. Belli et al (2020) observed that international collaboration in coronavirus research has been growing rapidly across the countries in 2020, when compared with the last two decades (2001–2019).

The thicker lines between USA~UK and USA~China are indicative of stronger linkage and their close association. Thus it identifies the closely associated countries of collaboration, as shown in Fig. 4. Deng et al (2020) also viewed that USA, China and UK are the three prominent players of international cooperation in the human coronavirus research. In a similar way, peer-group of institutions may be identified within a country. The collaboration network across many institutions of different countries is a valuable indicator of internationality for research and innovation (Perneger & Hudelson 2007).

Finding core-journals

This study looked into the sources of information (neo-micro-thoughts) available through journals preferred by researchers. Preferred journals are determined by the frequency of relevant articles published in the source. The highest number of articles about nCov-2 or COVID-19 are published in the BMJ Clinical Research Edition (266 articles) followed by Journal of Medical Virology (243), British Medical Journal (201), The Lancet (195), Nature (127), Journal of the American Medical Association (122), Journal of Infection (118), Lancet Infectious Diseases (117), Infection Control & Hospital Epidemiology (107), Clinical Infectious

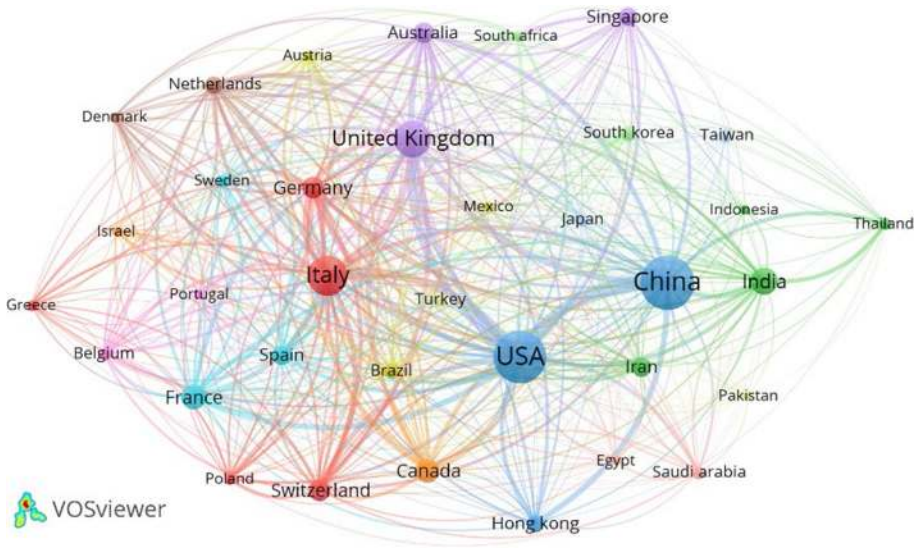


Fig. 4 Co-authorship network among the collaborating countries (top-twenty considered)

Diseases (100), Travel Medicine & Infectious Disease (100), etc. It implies that these journals were highly preferred by scientists to report their research findings.

In reverse, the journals that are often preferred by the researchers to grasp and channelize their scientific ideas – seem to be the core-journals. In ranking core journals, a new measure combining both ‘Impact Score of a journal’ (viz. Impact Factor or CiteScore) and ‘Frequency of relevant publications’ has been employed. Such a metrics can provide more rational and trusted view to assess the value of a source-title through weighted score, which may be utilized to identify potential journals with finer tunings. Thus, Weighted Score (WS) for each journal has been computed using the formula; $WS = \text{Impact Score (IS)} \times \text{Frequency (Freq)}$.

However, the Frequency of relevant publications when canonically combines with the average Impact Score i.e. the mean value of CiteScore and Impact Factor (Garfield, 2006) then optimally it could offer a normalized value for identifying core-journals (as presented in Table 4).

$$WS_{\text{Journal}} = \text{IS} \times \text{Freq} = \left\{ \frac{(\text{IF} + \text{CS})}{2} \right\} \times \text{Freq}$$

$$WS_{\text{Lancet}} = \left\{ \frac{(60.39 + 73.40)}{2} \right\} \times 195$$

$$= (66.90 \times 195) = 13045.50$$

Table 4 A composite measure of ranking core-journals (based on weighted score)

Rank	Journal name	IF-2019	CiteScore 2019	Avg-IS 2019	Freq. Publs	Weighted score
1	The Lancet (0140–6736, Elsevier)	60.392	73.4	66.90	195	13,045.50
2	Nature (0028–0836, Springer)	42.778	51.0	46.89	127	5955.03
3	BMJ Clinical Res. Ed. (0959–8146, BMJ Group)	30.223	6.5	18.36	266	4883.76
4	Journal of the Am. Med. Asso.(0098–7484, AMA)	45.540	26.3	35.92	122	4382.24
5	The British Med. Journal (0959–8138, BMJ Group)	30.223	6.5	18.36	201	3690.36
6	Lancet Infectious Diseases (1473–3099, Elsevier)	24.446	32.4	28.42	117	3325.14
7	Clinical Infectious Diseases (1058–4838, OUP)	8.310	12.5	10.41	100	1041.00
8	Journal of Infection (0163–4453, Elsevier)	4.842	8.0	6.42	118	757.56
9	Journal of Med. Virology (0146–6615, Wiley-BW)	2.021	4.0	3.01	243	731.43
10	Travel Med. & Invec.Disease (1477–8939, Elsevier)	4.589	4.7	4.64	100	464.00
11	Infection Control & Hosp. Epi. (0899-823X, CUP)	2.930	4.5	3.72	107	398.04

Frequency based ranking of the above journals has been changed substantially

So the potentially viable journals are identified using a combined score. These are *Lancet* (weighted score 13,045), *Nature* (5955), *BMJ Clinical Research Edition* (4884), *Journal of the American Medical Association* (4382), *British Medical Journal* (3690), *Lancet Infectious Diseases* (3325), and others. This active set of journals can be helpful to librarians to formulate a strategic plan for collection development. The finding also reveals that most of the preferred journals are published by Elsevier Publishing Group. So the researchers have maintained an international appeal in publishing their scientific results.

Citation-map of the publications

Scientific visualization typically relates to co-researchers through citation behavior. The phenomenon of citation denotes a “symbolic payment for an intellectual debt” (Small, 2004). It could be utilized to retrieve the authoritative sources of information by recognizing the visibility of the researchers. Citation count is, however, a convenient way of measuring the influence (quality) of research publications. Many scholars believe that citation indicators can be more effective than peer-judgment for evaluating interdisciplinary research (Rafols et al., 2012). So an attempt has been made to evaluate the influence of the publications through citation counts, using Scopus as on 31 May 2020.

Within a few months, the publication dataset (14,588) has received a total of 58,066 citations, which are distributed in Table 5. It appeared that the top 1.2% (175) articles occupied almost 50% (29,079) of the citations, and top-1000 (7%) papers together received 80% of the total score. An average of 4 citations per paper reasonably justifies the influence of the publication dataset. However, the mean number of citations per cited publication ($10.8 \approx 11$) is found very impressive. On the contrary, about 9213 (63%) of the publications are yet to receive a citation.

The h-index value is derived as 92. It means that 92 publications have received a minimum of 92 citations each. The g-index is computed as 170, implies that top-170 papers received a total of (170×170) 28,900 citations. A sum of 1010 publications received at least 10 citations each (i10-index).

Table 5 Distribution of citations into various classes (of different ranges)

Period	Range of citations	Publication Freq	Cu. Freq	Cu. %	Total citations	Cu. citations	Cu. %
2020 (January–May)	1001–2000	03	3	0.02	3838	3838	6.61
	501–1000	08	11	0.08	5771	9609	16.55
	201–500	22	33	0.23	6201	15,810	27.23
	101–200	47	80	0.55	6317	22,127	38.11
	51–100	111	191	1.31	7781	29,908	51.51
	21–50	328	519	3.56	10,288	40,196	69.22
	11–20	403	922	6.32	5770	45,966	79.16
	5–10	828	1750	12.00	5738	51,704	89.04
	1–4	3625	5375	36.85	6362	58,066	100.00
	00	9213	14,588	100.00	0000	58,066	100.00

Top-cited 175 publications together received 29,079 (almost 50%) citations

The highly-cited (top-ten) publications that received a minimum of 500 citations are shown in Table 6. Enlisted publications are considered as the most influential contributions of the new coronavirus (nCoV-2) research. Notably, all those are open access publications and published in high-impact journals like *The Lancet* (IF=59.1), *JAMA* (51.3), *Nature* (43.07), and *New England Journal of Medicine*.

Concluding remarks

The study reveals that global scientific communities are proactively fighting with this most extended sequence of RNA virus (Majumder, 2020). Newer insights are being made regularly to combat the pandemic situation around the globe. The World Health Organization has imposed stringent guidelines to protect the dynamic transmissions of COVID-19. Global researchers are equally prompting the shreds of evidence through empirical results of their research and practices for social welfare. Consequently, a vast amount of scholarly publications about nCoV-2 and COVID-19 has occurred during a few months in 2020. The publications are still booming to achieve success through the combined efforts of the researchers worldwide. Such a proliferation of scientific information is unprecedented in the history of science. Thus it signifies the importance of this evaluative study by recognizing the researchers' accountability at the time of medical emergency on Earth.

Findings reveal that the growth of coronavirus research has followed a graph of multiple peaks and valleys, where two peaks are very significant. The first peak appeared in 2003 (with 577% growth rate) when the SARS-CoV-1 epidemic contaminated over 28 countries. The next surge of the publication growth (1600%) is indicating a knowledge outburst. The study found that a vast majority (about 80%) of the publications are freely accessible (via Gold Open Access), and most of the contributions were made in the English language (93%). Worthy of mentioning, the information contents have published in more than 500 journals from 240 academic publishers. Elsevier publishing group occupies the central position, followed by NLM-Medline, Wiley-Blackwell, Springer-Nature, Lancet-Group, BMJ-Group, WB~Saunders, Taylor & Francis, Nature Research, American Medical Association, CUP, Academic Press, and BioMed Central.

While the top-ten publishers occupied almost 70% of the scientific contents, about 25% of the researches were granted by (approx) 300 funding agencies. Among the significant funding bodies – National Natural Science Foundation of China, National Institutes of Health (USA), National Basic Research Program of China (Program 973), National Institute of Allergy and Infectious Diseases (USA), Wellcome Trust of London (UK), National Institute for Health Research (UK), Pfizer Ltd., Bill & Melinda Gates Foundation, National Science Foundation are prominent. The result shows that majority of the contributions have occurred in Medical Science, focusing on virology, immunology, epidemiology, pharmacology, public health, critical care, and emergency medicine. However, the co-occurrence map of the author-generated keywords has shown the closely associated terms like virus transmission, infection control, asymptomatic, quarantine, pneumonia, disease severity, clinical trials, viral pathogenesis, pandemic, infection risk, mortality, and others. These are the active topics of research.

Findings also reveal that there is a drastic variation of research priorities amongst the countries. Academic hubs of the new coronavirus research are located mostly in the USA and China, followed by Italy, UK, India, France, Canada, and Germany. All these countries are affected severely by COVID-19. The G7 countries together produced 50%

Table 6 Ranked list of highly-cited publications of the new coronavirus research

SI	Publication details	Inl. IF	Citations
1	Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., ... & Cheng, Z. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. <i>The Lancet</i> , 395(10,223), 497–506. {Open Access}	(59.1)	1765
2	Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., ... & Zhao, Y. (2020). Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan. <i>Journal of the American Medical Association (JAMA)</i> , 323(11), 1061–1069. {Open Access}	(51.3)	1071
3	Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., ... & Niu, P. (2020). A novel coronavirus from patients with pneumonia in China, 2019. <i>New England Journal of Medicine</i> , 382(8), 727–733. {Open Access}	(70.67)	1002
4	Guan, W. J., Ni, Z. Y., Hu, Y., Liang, W. H., Ou, C. Q., He, J. X., ... & Du, B. (2020). Clinical characteristics of coronavirus disease 2019 in China. <i>New England Journal of Medicine</i> , 382(18), 1708–1720. {Open Access}	(70.67)	980
5	Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., ... & Yu, T. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. <i>The Lancet</i> , 395(10,223), 507–513. {Open Access}	(59.1)	947
6	Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., ... & Xing, X. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. <i>New England Journal of Medicine</i> , 382(13), 1199–1207. {Open Access}	(70.67)	820
7	Wu, Z., & McGoogan, J. M. (2020). Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. <i>Journal of the American Medical Association (JAMA)</i> , 323(13), 1239–1242. {Restricted}	(51.3)	639
8	Zhou, P., Yang, X. L., Wang, X. G., Hu, B., Zhang, L., Zhang, W., ... & Chen, H. D. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. <i>Nature</i> , 579(7798), 270–273. {Open Access}	(43.07)	634
9	Chan, J. F. W., Yuan, S., Kok, K. H., To, K. K. W., Chu, H., Yang, J., ... & Tsoi, H. W. (2020). A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. <i>The Lancet</i> , 395(10,223), 514–523. {Open Access}	(59.1)	623
10	Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., ... & Guan, L. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. <i>The Lancet</i> , 395(10,229), 1054–1062. {Open Access}	(59.1)	595
11	Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., ... & Bi, Y. (2020). Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. <i>The Lancet</i> , 395(10,224), 565–574. {Open Access}	(59.1)	533

of the global research output on nCoV-2. On the contrary, most of the African countries have produced very little. A few institutions have shown their phenomenal activity in the COVID-19 research. Among the productive institutions; Huazhong University of S&T (China), Tongji Medical College at Wuhan (China), Harvard Medical School at Boston (USA), University of Milan (Italy), INSERM at Paris (France), University College of London (UK), University of Toronto (Canada), University of Oxford (UK), University of Washington (USA) are leading.

The study noted an encouraging trend of collaborative research across different countries and disciplines. About 80% of publications are collaborative with a good number of mega-authorship publications (13%); where the values of Collaborative Index (6.46), Degree of Collaboration (0.79), and Collaborative Coefficient (0.59) are very much indicative. The co-authorship network of the collaborating countries has shown the geographical diversity and linkages of the collaborating authors. Visualization map indicates that profound collaborators are located mostly in the USA, UK, Italy, China, and Australia. It also identified the potentially viable (core) journals of the new coronavirus research using a combined metrics. These are Lancet (weighted score 13,045), Nature (5955), BMJ Clinical Research Edition (4884), Journal of the American Medical Association (4382), British Medical Journal (3690), Lancet Infectious Diseases (3325) among the others.

Finally, it analyzed the impact (quality) of the publications through citation counts. The publication dataset has received a total of 58,066 citations within a couple of months. The mean number of citations per cited publication ($10.8 \approx 11$) is fairly impressive. The values of h-index (92), g-index (170), and i10-index (1010) are also signaling for the quality. Top 175 articles occupied almost 50% of the total citation score, and top-ten highly cited contributions received a minimum of 500 citations each. All these articles (except one) have published without any pay-wall (i.e. open-access) and also appeared in the high-impact journals like Lancet (IF=59.1), JAMA (51.3), New England Journal of Medicine (70.67), and others.

These results are indicative for recognizing the accountability of the researchers to defeat this global crisis. Indeed this visualization of the publications about COVID-19 that were channelized until June 2020 can be seen as a first-hand result based on a primary set of indicators. It is, however, necessary to visualize the knowledge outburst more exhaustively in different dimensions. Therefore, a wider study could be taken up by conceptualizing the empirical results at the micro-level.

The modeling for a knowledge graph of this synergic response could be the future direction of research. Besides that, one can make use of Almetrics (an alternative to traditional citation-metrics) to monitor attention surrounding the publications of this newly configured area of research.

The above findings shall be helpful to propose future strategies in the research of this emerging area. We believe that the superfluous growth of COVID research has to shed light on the pathway of developing a vaccine of the new coronavirus disease.

Acknowledgement The author gratefully acknowledge with thanks to anonymous reviewers for their valuable comments and fruitful opinions on a draft of this paper. He also wishes to thank his beloved son Adrito for providing constant inspiration to prepare this paper during the lockdown period.

Compliance with ethical standards

Conflict of interest There is no conflicts of interest.

References

- Ajiferuke, I., Burell, Q., & Tague, J. (1988). Collaborative coefficient: A single measure of the degree of collaboration in research. *Scientometrics*, *14*(5–6), 421–433.
- Belli, S., Mugnaini, R., Baltà, J., & Abadal, E. (2020). Coronavirus mapping in scientific publications: when science advances rapidly and collectively, is access to this knowledge open to society? *Scienometrics*, *124*, 2661–2685. <https://doi.org/10.1007/s11192-020-03590-7>.
- Chahrouh, M., Assi, S., Bejjani, M., Nasrallah, A. A., Salhab, H., Fares, M., & Khachfe, H. H. (2020). A bibliometric analysis of Covid-19 research activity: A call for increased output. *Cureus*. <https://doi.org/10.7759/cureus.7357>.
- Dehghanbanadaki, H., Seif, F., Vahidi, Y., Razi, F., Hashemi, E., Khoshmirsafa, M., & Aazami, H. (2020). Bibliometric analysis of global scientific research on Coronavirus (COVID-19). *Medical Journal of The Islamic Republic of Iran (MJIRI)*, *34*(1), 354–362.
- Deng, Z., Chen, J., & Wang, T. (2020). Bibliometric and visualization analysis of human coronaviruses: prospects and implications for COVID-19 research. *Frontiers in Cellular and Infection Microbiology*, *10*(581404), 529–535. <https://doi.org/10.3389/fcimb.2020.581404>.
- Feller, I., Glasmeier, A., & Mark, M. (1996). Issues and perspectives on evaluating manufacturing modernization programs. *Research Policy*, *25*(2), 309–319.
- Garfield, Eugene (2006). The history and meaning of the Journal Impact Factor. *Journal of the American Medical Association*, *295*(1), 90–93. [Journal Impact Factor (JIF) is a bi-product of the Web of Science (JCR) Citation Database produced by the Clarivate Analytics. URL: <https://clarivate.com/webof-sciencegroup/solutions/journal-citation-reports/#>
- Haghani, M., & Bliemer, M. C. J. (2020). Covid-19 pandemic and the unprecedented mobilisation of scholarly efforts prompted by a health crisis: scientometric comparisons across SARS, MERS and 2019-nCoV literature. *Scientometrics*, *125*, 2695–2726. <https://doi.org/10.1007/s11192-020-03706-z>.
- Hossain, Md. M. (2020). Current status of global research on novel coronavirus disease (Covid-19): A bibliometric analysis and knowledge mapping. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3547824>.
- Lawani, S. M. (1980). *Quality, collaboration and citations in cancer research: a bibliometric study (Doctoral dissertation)*. USA: Florida State University.
- Lou, J., Tian, S. J., Niu, S. M., Kang, X. Q., Lian, H. X., Zhang, L. X., & Zhang, J. J. (2020). Coronavirus disease 2019: a bibliometric analysis and review. *Eur Rev Med Pharmacol Sci*, *24*(6), 3411–3421. https://doi.org/10.26355/eurrev_202003_20712.
- Majumder, Partha P. (2020). Lockdown maths: why we must comply (guest column). *The Telegraph (online edition)*, 25 March (Sunday), e-paper.
- Nasab, F. R., & Rahim, F. (2020). Bibliometric analysis of global scientific research on SARSCoV-2 (COVID-19). *medRxiv*. <https://doi.org/10.1101/2020.03.19.20038752>.
- Pal, J. K. (2015). Scientometric dimensions of cryptographic research. *Scientometrics*, *105*(1), 179–202. <https://doi.org/10.1007/s11192-015-1661-z>.
- Perneger, T. V., & Hudelson, P. M. (2007). How international is the International Journal for Quality in Health Care? *International Journal for Quality in Health Care*, *19*(6), 329–333.
- Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., & Stirling, A. (2012). How journal rankings can suppress interdisciplinary research: A comparison between innovation studies and business & management. *Research Policy*, *41*(7), 1262–1282.
- Small, H. G. (2004). On the shoulders of Robert Merton: towards a normative theory of citation. *Scientometrics*, *60*(1), 71–79.
- Subramanyam, K. (1983). Bibliometric studies of research in collaboration: a review. *Journal of Information Science*, *6*(1), 33–38.
- Teixeira da Silva, J. A., Tsigaris, P., & Erfanmanesh, M. (2020). Publishing volumes in major databases related to Covid-19 (Letter to the Editor). *Scientometrics*. <https://doi.org/10.1007/s11192-020-03675-3>.
- VOSviewer (Software for Visualizing Scientific Landscapes). Developed by Centre for Science and Technology Studies, Leiden University, The Netherlands. Retrieved from <http://www.vosviewer.com/> (accessed on 31 May 2020).
- Yu, Y., Li, Y., Zhang, Z., Gu, Z., Zhong, H., Zha, Q., et al. (2020). A bibliometric analysis using VOSviewer of publications on COVID-19. *Annals of Translational Medicine*, *8*(13), 816–822. <https://doi.org/10.21037/atm-20-4235>.
- Zhang, L., Zhao, W., Sun, B., Huang, Y., & Glanzel, W. (2020). How scientific research reacts to international public health emergencies: a global analysis of response patterns. *Scientometrics*, *124*, 747–773. <https://doi.org/10.1007/s11192-020-03531-4>.

Authors and Affiliations

Jiban K. Pal¹ 

✉ Jiban K. Pal
jiban@isical.ac.in

¹ Library, Documentation and Information Science Division, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India