Journal of Scientometric Res. 2022; 11(1):89-99
http://www.jscires.org

Research Article

Intellectual Structure Mapping of Sickle Cell Anemia Research in India: A Scientometric Analysis

Dhruba Jyoti Borgohain[1], Manoj Kumar Verma[1,*], Brady Lund[2]
[1]Department of Library and Information Science, Mizoram University, Aizawl, Mizoram, INDIA.
[2]School of Library and Information Management, Emporia State University, KS, US.

ABSTRACT
There are limited studies which focused on network analysis and visualization of Indian Research performance in Sickle Cell Anaemia. This study is an attempt to bridge this gap using metrics with a view to understand the present status of research at the global, national, institutional, author and source level. The study is based on publication and citation data sourced from Scopus during 1958 to 2020. The bibliographic data was statistically analyzed on various metrics such as country collaboration, document type, productive author, journal, highly cited works, productive year, author affiliation. The U.S.A, U.K. and India are highly collaborating in research on "Sickle Cell Anemia". Canada, the USA and Jamaica are highly cited nations. Indian Journal of Hematology and Blood Transfusion is the most productive journal. In addition, the study also investigates and maps productive institutions, collaborations among these institutions, key authors, key source journals and also most significant keywords in the subject thereby visually presenting their inter-relationships using Biblioshiny and VOSviewer software. Results and findings from this study describes the progress made by India through research on this deadly genetic disorder and the future scope as well as trends which will be very useful for researcher working and also having zeal or enthusiasm in this area.
Keywords: Sickle Cell Anemia, RBCs, Genetic Disorder, India, Scientometrics, Bibliometrics, Scopus.

Correspondence
Manoj Kumar Verma
Department of Library and Information Science, Mizoram University, Aizawl-796004, Mizoram, INDIA.
Email id: manojdlis@mzu.edu.in
ORCID ID: 0000-0002-3009-3258

Received: 30-09-2021;
Revised: 09-03-2022;
Accepted: 19-03-2022.
DOI: 10.5530/jscires.11.1.9

INTRODUCTION

Sickle cell Disease (SCD) is a class of genetic blood disorder meaning inherited from parents. The area considered under study i.e., "Sickle Cell Anemia" (SCA) is a type of Sickle Cell Disease. This disease has been reported first in 1846 during an autopsy of an executed slave was discussed and the finding of the study was absence of Spleen in his body. Now, spleen is an organ in the left part of the abdomen protected by the rib cage.[1] This generally functions as a blood filter in all vertebrates and particularly it recognizes old or damaged Red Blood Cells (RBCs) and eliminates them from our body by breaking them down and storing useful components like iron in the process. This in turn results in circulation of clean blood in our body and functioning of blood at its best. Now, a person with SCA will have abnormality in hemoglobin, the oxygen carrying protein found in RBCs leading to a rigid sickle like shape of these cells. This problem due to this disease can be observed at an early age of 5 to 6 months via number of health problems like attack of pain, acute anemia, swelling of body parts like hands and feet, bacterial infection and also stroke. The body organs of person suffering from this ailment will cease to function gradually due to limited oxygen supply because of gradual decrease in the number RBCs which will limit to null in course of time.[2] This disease is deadly because it is inherited implying that a generation with SCA will pass on this curse to the upcoming progenies limiting their life expectancy within 40 to 60 years. The treatment of this disease needs highly sophisticated and state-of-the-art medical equipment's and tools because this involves works in cellular level. Research on treatment procedures resulted in evolution of three techniques viz. (i)Umbilical Cord Blood Transplant which needs suitable donor and reported to be suitable in only 10% of people, (ii) Gene Therapy uses normal copy of the genes that is mutated (iii)Hematopoietic Stem Cell Transplantation method has no evidence of treating people suffering from SCD.[3]

In Indian context, this disease is found to be common in some ethnic groups of Central India where the presence of this genetic disorder has raised from 9.4% to 22.2% in regularly detected areas of Madhya Pradesh, Rajasthan and Chhattisgarh.[4] A study on mortality in Sickle Cell Disease during 2008 to 2018 in an abnormal community in the Gudalur Valley, Nilgiris, Tamil Nadu, India.[5] 157 patients taken as sample who belongs to Paniya, Betta Kurumba, Kattunyakan and Mullu Kurumba tribes. The study reveals that during the study period there were 22 deaths and all are from the Paniya

tribe. Twelve deaths (54.5%) occurred in the hospital and the remaining at home (45.5%), reflecting a crude mortality rate of 140 per 1000 population. Twenty-five percent of deaths occurred in the 6–18 age group. There were no deaths in the 0–5 age group. The median age of death was 25 years, which was 30 years less than in the non-SCD aboriginal population. The leading causes of death were acute chest syndrome, anaemia, and sepsis among the SCD patients and stroke and suicides in the non-SCD aboriginal population. Given the brutality of this genetic ailment in India as well as in global scenario it is considered worthwhile that a scientometric assessment be undertaken to understand and describe the current state of research on "Sickle Cell Anemia" based on Indian literatures published and indexed in Scopus till date.

Literature Review

Over the past several decades, quite a few studies have been undertaken on different human ailments be it genetic or viral using bibliometric methods. Some of the works are reviewed as follows. A very high impactful analysed the pattern of literature growth, global publication share, ranking, authorship pattern, collaborative coefficient, productivity and impact of most productive institutions and authors, highly cited articles based on data obtained from Scopus database on Chronic Liver disease (CLD) in SAARC countries.[6] The study reveals that the SAARC nations contributed 2312 documents during 1996 to 2015 which is only 3.49% of the global output of 66200 publications. The also reflects that the amount of literature has considerably increased over the last five years of the period of study. India is leading among the SAARC nations. A bibliometric analysis of *Plesiomonas* related research with data from Web of Science during 1990 to 2017 reveals that a total 155 articles were published in the survey period with annual growth of 0.8%. The USA ranks first in terms of number of articles (n=29) and total citations (451).[7] The research collaboration was also low with collaboration index of 3.32. This bibliometric analysis reveals that there is a global diminishing research in Plesiomonas and greater research outcome is from high income countries compared to others. Another bibliometric analysis on prediction of infectious disease with data sourced from Web of Science.[8] The 1880 documents published on the topic has been analyzed and this analysis reveals that the publications were published in 427 different journals with 11 different document types and most common is articles (1618). The study reveals that *Nature* Journal is the top cited journal with 781 citations, followed by *PLoS ONE* with 707 citations. The USA is the most productive nation with 749 documents. A study investigated Coronavirus literature using bibliometric indicators taking data from 1970 to 2019 with data sourced from Scopus.[9] The study was carried out using the keyword Coronavirus and analysed for annual growth, productive countries, institutes, authors, journals, highly cited papers, and research focus using keywords. Most of the research publications were from the USA (31.67%), and the University of Hong Kong was the most productive institute. A study evaluated the global research output (820 records) on "Use of Convalescent Plasma Therapy for COVID-19" on metrics with the aim to understand the current status of research at the global, national, institutional, and individual author level.[10] The study is based on publications and citations data sourced from the Scopus database during 2020-21. The publications and citations data were statistically analysed on various metrics such as document type, country of publication, collaboration patterns, author affiliation, journal name, and citation patterns.

Recently, a few attempts have been made to analyze literature on Sickle Cell Anemia. Those are reviewed as follows. A study analysed the literature on Sickle Cell Anaemia in Nigeria with data extracted from PubMed listed during 2006 and 2016.[11] *Nigerian Journal of Clinical Practice* is the most productive journal from the nation. The highest number of contributions are with the USA followed by Italy. Another study analyzed the global literature on Sickle Cell Anemia using bibliometric indicators with literature published during 1997 to 2017 indexed in Scopus database.[12] 19,921 publications were recovered during the period where majority are journal articles. The findings reveal that *Blood* is the highly productive source, the USA is the leading nation and India is on the 5th position in global perspective. Keywords like human, hydroxyurea, blood transfusion, controlled study, clinical research, anemia, pathogenesis are most common. There came another study on Sickle cell disease in global perspective with data extracted from Web of Science indexed during 1900 to 2020 applying the bibliometric indicators.[13] This study also reveals that *Blood* is the most productive journal and the most productive author is Sergeant, GR. The most prominent keywords are anemia, children, disease and management. The USA is the most productive nation.

Bibliometrics applies mathematical and statistical methods to brief scientific activities in a subject leading to help in identifying research frontiers, trending areas and rising patterns based on literatures from various relevant databases.[14-15] Utilization of several visualization tools like VOSviewer,[16] Biblioshiny,[17] CiteSpace,[18] HistCite[19] to develop knowledge and network maps, analyze latest research progress and visualize the trends and co-authorships in scientific publications.[20-21] This study attempts to prepare network visualization maps for different bibliometric parameters using the analysis units like journals, authors, organizations, countries and keywords.

No one can deny that a lot of bibliometric analysis are available on topics related to diseases. But not a single study has been found to be published in a Scopus or Web of Science indexed

journal on Indian research output in Sickle Cell Anemia with data extracted from Scopus that too from 1958 till date. So, this study attempts to fulfill this research gap.

The study attempts to describe the Indian contribution on this disease using metrics (i) the extent of global research collaboration of India (ii) analyzes the author's contribution on the basis of number of publications and citation impact (iii) investigate the most productive institutions (iii) identify the highly cited works and the document forms (iv) to analyze the network of country, author, organizational co-authorship and bibliographic coupling of journals (v) analyze the cluster and network of keywords co-occurrence hence evaluate the mostly occurring keywords on the basis of frequencies, mapping word growth and map trending topics in Sickle Cell Anemia Research. The results of this study will be relevant to researchers, physicians and health policy makers as well as government.

METHODS

Identification of the Search Strategy

A well- defined search strategy was used to retrieve and download publications data from the Scopus database. The search for global literature published on "Sickle Cell Anemia" was conducted with no starting date specified. A suitable search strategy for retrieving data from the Scopus database was developed so as to have a reliable set of data for analysis and to obtain accurate and precise results. An "All field" search has been conducted with search terms related to "Sickle Cell Anemia" as ((ALL ("Sickle Cell Anemia" OR "Anemia" OR "Genetic Blood Disorder") OR (TITLE ("Sickle Cell Anemia" OR "Anemia" OR "Genetic Blood Disorder") AND (LIMIT-TO (AFFILCOUNTRY, "India")))). Here, the search operator "OR" is used instead of "AND" operator because use of "OR" operator gives more broader search results making the findings of the study more efficient, precise and reliable. The database search resulted 57,436 results as global output. Since we are concerned with India, so the data is next limited to India in the country box and this resulted 1443 literature from the initial year of indexing 1958 to 2020. The Indian output was subsequently analyzed for publication output by author, affiliation, journal, country of publication, top cited countries, top cited documents, year-wise productivity and prominent keywords are analyzed to predict the trending areas of research in the field.

Analysis Tools and Techniques

Analysis and tabulation of data are done using MS-Excel. For mapping the data Biblioshiny and VOSviewer are used. The indicators analyzed in the study as per the objectives are top collaborative countries with India on the basis of number of publications and citations; productive authors; productive institutions or organizations on the basis of number of publications, total local and global citations; highly cited research work on the basis of local and global citations, most common medium of communication on the basis of number of records; most productive and highly cited journal; most productive year, significant keywords and trending topics. The analyzed and mapped data is depicted in tables, network visualization maps and interpreted objective-wise in the following segment. In particular, Figure 1 is generated using the trending bibliometric data analysis tool Biblioshiny Web interface in R-studio while Figure 2 is generated using the popular network visualization and mapping tool VOSviewer. If we see the former has depicted the authors on the basis of productivity and the later represents a network visualization of highly cited authors. The default normalization method for the analysis is "Association Strength" for creating Figures 2-6.

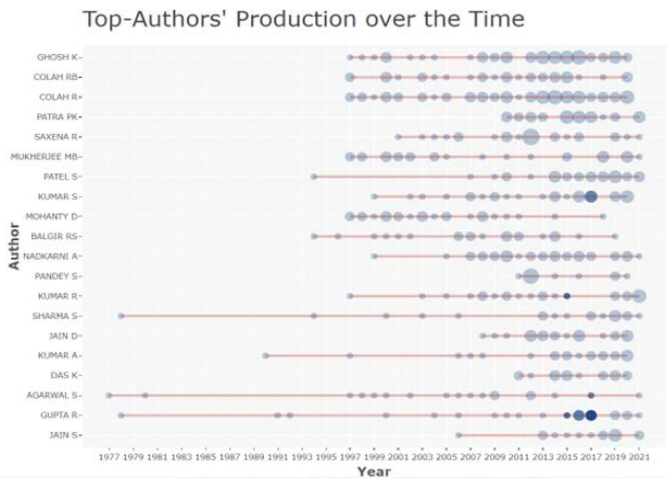

Figure 1: Mapping of most Productive Authors (blue circles showing number of publications).

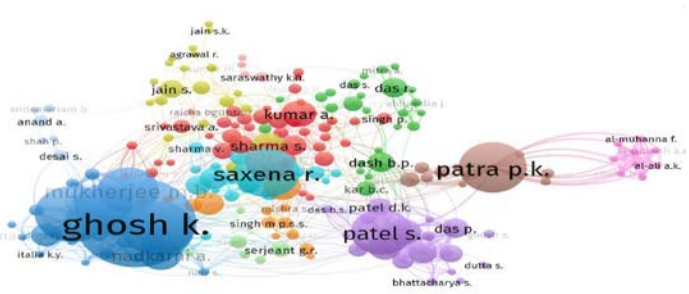

Figure 2: Visualization of author co-authorship.

RESULTS

Analysis and tabulation of data are done using MS-Excel. For mapping the data Biblioshiny and VOSviewer are used. The indicators analyzed in the study as per the objectives are top collaborative countries with India on the basis of number of publications and citations; productive authors; productive

institutions or organizations on the basis of number of publications, total local and global citations; highly cited research work on the basis of local and global citations, most common medium of communication on the basis of number of records; most productive and highly cited journal; most productive year, significant keywords and trending topics. The analyzed and mapped data is depicted in tables, network visualization maps and interpreted objective-wise in the following segment.

Country co-authorship

In all, 120 countries have participated in global research on "Sickle Cell Anemia" but the distribution of collaborative productivity by these nations is not normal more specifically it is highly skewed.

A country co-authorship map is generated using the popular network visualization and mapping tool VOSviewer[14] [Figure 3]. It provides a visual presentation of comparative productivity and the nature of the collaboration in research within the field of study. Taking minimum of 2 countries which are co-authoring as threshold, 50 countries are found to be networked. The links shown using straight lines depicts the collaboration between the countries. The thickness of links between the countries represents the strength of their collaboration. Overall, these 50 countries are divided into 7 clusters. Cluster 1 (Red) has 18 countries, some of them are Germany, Switzerland, United Kingdom, Spain, Thailand. Cluster 2 (Green) has 12 countries in total and some of them are China, Italy, Iran, Saudi Arabia, Turkey, Bulgaria. This is followed by Cluster 3 (Deep Blue) with six countries, these are Canada, Finland, Hong Kong, Nigeria, Singapore, South Africa. Cluster 4 (light yellow) has 5 countries, Bangladesh, Japan, Pakistan, South Korea, Taiwan. Cluster 5 (Purple) has 4 countries including India and Cameroon, Jamaica, Libyan Arab Jamahiriya. Cluster 6 (Shallow Blue) has 3 countries, Australia, Sri Lanka and United Arab Emirates. Lastly, cluster 7 (Orange) has 2 countries United States and Russian Federation.

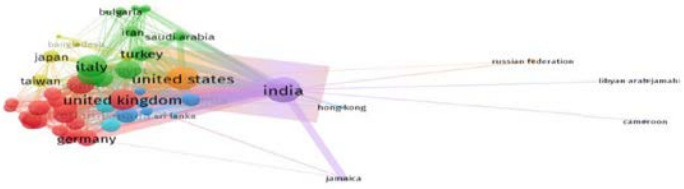

Figure 3: Mapping of country co-authorship.

Table 1: Top 10 Productive Author on the basis of publication and citation impact.

Sl. No.	Publication Impact				Citation Impact			
	Author	Affiliation	Recs	Citations	Author	Affiliation	Recs	Citations
1	Colah B. Roshan	National Institute of Immunohematology (ICMR), Mumbai	67	1010	Gupta Renu	All India Institute of Medical Sciences, New Delhi	12	11664
2	Ghosh Kanjaksha	National Institute of Immunohematology (ICMR), Mumbai	55	657	Basu Swati	Dept of Microbiology, School of Tropical Medicine, Kolkata	9	9748
3	Patra Pradeep Kumar	Pt. J.N.M. Medical College, Raipur	32	192	Banerjee Anuradha	Jawaharlal Nehru University	9	9706
4	Saxena Richa	Dept of Reproductive Biology, ICMR, NewDelhi	30	261	Sharma Satish	Mahatma Gandhi Institute of Medical	14	7362
5	Mukherjee Malay B	National Institute of Immunohematology (ICMR), Mumbai	28	395	Gupta Vineeta	Institute of Medical Sciences, B.H. U	13	7173
6	Patel Jyotish	Vision Medical Foundation for Rural Health and Research, Gujrat	28	150	Kumar Preetham	Manipal University, Karnataka	14	4284
7	Mohanty Dipika	National Institute of Immunohematology (ICMR), Mumbai	26	504	Kumar Sunil	Department of Clinical Pathology, Lok Nayak Hospital, New Delhi	25	3048
8	Kumar Sunil	Department of Clinical Pathology, Lok Nayak Hospital, New Delhi	25	3048	Ghosh Kanjaksha	National Institute of Immunohematology (ICMR), Mumbai	55	657
9	Balgir Ranbir S	Regional Medical Research Centre (ICMR), Bhubaneswar	22	339	Colah B. Roshan	National Institute of Immunohematology (ICMR), Mumbai	34	586
10	Nadkarni Anita H	National Institute of Immunohematology (ICMR), Mumbai	19	214	Mohanty Dipika	National Institute of Immunohematology (ICMR), Mumbai	26	504

Productive Authors

In total, 320 authors have participated globally in "Sickle Cell Anemia" Research with some authors contributing more articles in this area than others. Table 1 shows top 10 author's scientometric profile based on publication and citation. First three columns in Table 1 shows prolific authors on the publication impact and the other three columns listed authors with citation impact. The results in Table 1 implies that an author who is highly productive (with maximum number of publications) may not be impactful (may depreciate in number of citations). Top publication output by Colah B. Roshan, Ghosh Kanjaksha of National Institute of Immunohematology (ICMR), Mumbai and Patra Pradeep Kumar of Pt. J.N.M. Medical College, Raipur with 67, 55 and 32 publications respectively. In terms of impact of the authors the highest number of citations are received by Gupta Renu of All India Institute of Medical Sciences, New Delhi followed by Basu Swati, Dept of Microbiology, School of Tropical Medicine, Kolkata and Banerjee Anuradha with 11664, 9748 and 9706 citations respectively. Three of them have 12, 9 and 9 publications respectively.

In Figure 2, we can observe that some authors have consistently produced research work like Ghosh Kanjaksha and some are very productive earlier but vanishes later on like Agarwal Sarita. The radius of the circles is directly proportional to the productivity of the author. The circles with greater radius for a respective year implies the author to be more productive in that year. In Figure 3, a network map of co-authorship of authors. The size of the circle and the font size of the letters naming the author is directly proportional to the number of documents of the authors. The colorful links between the authors indicates the co-authorship nature among the authors. Taking minimum number of documents of an author as 4, 208 authors are found to meet this threshold. For each 208 authors, the total link strength of the co-authorship links with other authors is calculated. The authors with greatest total link strength are selected by VOSviewer software to prepare the network map. From the 208 authors, 198 authors only form the largest set of connected authors. Hence, the network map will represent only these 198 authors. Now, these 198 authors are divided into 12 clusters, each cluster with different colors, on the basis of their closeness in research work. Cluster 1 (Red) has 41 authors in total, some of them are Kumar, A; Singh, P; Gupta, R; Srivastava, A. Cluster 2 (green) color has 25 authors, some of them are Kar, BC; Dash, BP; Das, R; Das S. Cluster 3 (Deep Blue) has 23 authors connected. Some of them are Ghosh, K; Mukherjee, MB; Nadkarni, A; Italia, KY; Colah, RB; Colah, R. Cluster 4 (yellow) has 21 authors, some are Jain, S; Agarwal, R; Jain, SK. Cluster 5 (violet) has 19 authors like Das, P; Patel, DK; Patel, S; Dutta, S; Bhattacharya, S. Cluster 6 (shallow blue) has 18 authors in club. Some in this cluster are Saxena, R; Gupta, A; Mahapatra, M; Choudhry, VP. Cluster 7 (Orange) has 12 authors like Singh, MPSS; Agarwal, S; Mishra, S. This is followed by cluster 8 (shallow brown) has 11 authors like Patra, PK; Patra, S; Singh, I; Verma, H. Cluster 9 (purple) has 9 authors in total. Some of them are Al-ali, AK; Al-muhanna, F; Al-rubaish, AM. Cluster 10 (shallow red) has 7 authors, some of them are Saraswathy, KN; Agarwal, A; Mitra, S; Tiwari, A. Cluster 11 (shallow green) has 6 authors in total, these are Serjeant, GR; Shah, A; Patel, P; Patel, J; Patel, A, Bhukhanvala, DS. Lastly, cluster 12 (light blue) has 6 authors and these are Anand, A; Andemariam, B; Desai, S; Javidi, B; Shah, P; Shah, S [Figure 3].

Organizational Productivity

In total, nearly 98 organizations participated in Indian "Sickle Cell Anemia" research and descriptive analysis of the dataset

Table 2: Most Productive Institution-wise records.

Sl. No.	Institutions	No. of Pub.	TLCS	TGCS
1.	All India Institute of Medical Sciences	318	7502	32754
2.	National Institute of Immunohaematology	155	5264	15491
3.	Government Medical College, Nagpur	96	1865	4478
4.	Not Reported	69	231	2167
5.	CSIR Institute of Genomics and Integrative Biology	92	994	3278
6.	Post Graduate Institute of Medical Education and Research	117	1378	4234
7.	Veer Surendra Sai Medical College	61	546	1489
8.	Christian Medical College	54	1180	4578
9.	Medical College	38	177	1355
10.	Inst. of Immunohematology (ICMR)	99	3785	11722
11.	Manipal University	32	704	1854
12.	Sanjay Gandhi Postgraduate Institute of Medical Sciences	32	2530	8453
13.	Sri Ganga Ram Hospital	32	1714	2864
14.	VIT University	29	751	1864
15.	Banaras Hindu University	27	1428	2568
16.	Maulana Azad Medical College	27	1244	2868
17.	Saha Institute of Nuclear Physics	27	1458	3210
18.	Jawaharlal Nehru University	26	1842	3356
19.	Jawaharlal Nehru Medical College	25	1241	2275
20.	University of Delhi	25	1947	2120

TLCS = Total Local Citation Score, TGCS = Total Global Citation Score

Figure 4: Organization Co-authorship.

implies abnormality in it. Table 2 provides a list of top institutes on the basis of publications and total citation output. As per productivity of organizations is concerned All India Institute of Medical Sciences (AIIMS) has highest number of 318 articles and 7502 local and 32754 global citations. AIIMS is followed by National Institute of Immunohematology with 155 total articles, 5264 and 15491 local and global citations respectively.

The collaboration network of authors is depicted in Figure 4 created using bibliographic network mapping software VOSviewer. The colored lines show the links between different organizations. The network map is prepared taking 1 as the minimum number of documents as the threshold and total 3347 organizations have minimum one document. Now, for the 3347 organizations, the total link strength of the co-authorship links with other organizations is calculated. Considering the number of documents as weight, the network map is created. Moreover, the font size of the names of the organizations and the size of the circles is proportional to the number of documents in the name of the organization.

34 organizations are connected and the network map for these organizations is prepared. These 34 organizations are further divided into 4 clusters, each cluster with separate colors. As depicted in Figure 7, organizations marked in red color are in Cluster 1 having 18 organizations. Some of them are Novartis Pharma Ag, Basel, Switzerland; Sickle Cell Disease Association of America, Baltimore; United States; Sickle Cell Society, London. Cluster 2 (green) has 8 organizations in club, some are Dr. NTR University of Health Sciences, Vijaywada, India; Institute of genetics and hospital for genetic diseases, Osmania University. In shallow blue color the cluster 3 has 5 organizations like Analysis group Inc., Montreal, Canada; National Minority Quality forum, Washington, DC, USA; Duke Adult Comprehensive Sickle Cell Centre, Durham, NC, United States; Novartis Pharmaceuticals Corporation, East Hanover, NJ, United States. Cluster 4 (shallow yellow) has 3 organizations namely Centre for Human Genetics, Central University of Punjab, Bhatinda, Punjab; Dr. NTR University of Health Sciences, Vijaywada, AP, India; Thalassemia and Sickle Cell Society, Hyderabad, India.

A three-field plot is used for further analysis which is created using Biblioshiny application [Figure 7]. The first field at the left are the countries, author is the second field at the middle and followed by affiliations in the third field at utmost right. The thickness of the rectangular boxes is proportional to the number of publications collaborated with the nations, of the authors and belonging to different affiliations. The curve lines starting from India touches most of the authors which implies their collaboration with Indian authors like Pandey, S; Sharma, S; Jain, S. But it is observed that these authors have also collaborated with the USA as some curved originating from the USA have also touched them like Sharma, S; Jain, S; Kumar, A. Authors like Gupta, R with huge number of publications affiliated to University of Washington, University of Oxford. Kumar, P to AIIMS and University of Washington. Kumar S has publications affiliated to CSIR Institute of Genomics and Integrative Biology. Pandey, S and Sharma, S have affiliations to Sanjay Gandhi Post Graduate Institute of Medical Science, Christian Medical College, Shri Ganga Ram Hospital. Kumar, S and Kumar, P with maximum collaboration with the USA, U.K., China, Australia and publication affiliations to AIIMS, CSIR Institute of Genomics and Integrative Biology, University of Washington, University of Oxford. Thickness of the box representing Gupta implies collaboration of the author with maximum countries and publications affiliated to maximum institutions.

Out of the total publications on "Sickle Cell Anemia" (1505 publications) in Indian perspective only 10 publications (0.65% of the total) accumulated highest citations, 180 to 30 since their publication in 1986. The distribution of top 10 highly cited papers is skewed: 2 papers received citations in the range 100–180, 3 papers are in average citation range of 50–100, 5 papers are in citation range of 1–50. Almost all highly cited papers are multi-authored. The articles in the citation range 100–180 are published in two journals in the list of top 10 highly articles. These are *Lancet* and the *Journal of the Association of Physicians in India*. The top cited work on Sickle Cell Anemia and the title of the work is Sickle Cell Disease in Orissa State authored by Kar, BC has citation count of 173 published in *The Lancet* followed by article titled Sickle Cell Disease in India by Mohanty, D which has about 94 citations published in *Current Opinion in Hematology* [Table 3].

Medium of Research Communication

In all, 1505 total research publications are encountered on "Sickle Cell Anemia Research" in Indian perspective as per data retrieved from Scopus. The publications are maximum in the form of journal articles. Of the total 1054 are journal articles (70.03%), 227 are review papers (15.08%), 87 letters (5.78%), 46 conference papers (3.05%), 43 book chapters (2.85%) which takes the majority share. The implication of this dataset is that the authors are more inclined to publish their research work in journals rather than other forms as these counts for more visibility, impact and durability.

Table 3: Top 10 Highly Cited Works.

Document	DOI	Year	Local Citations	Global Citations	Local Citations (%)
Kar BC, 1986, Lancet	10.1016/S0140-6736(86)92205-1	1986	49	124	39.52
Mohanty D, 2002, Curr Opin Hematol	10.1097/00062752-200203000-00006	2002	39	55	70.91
Kar BC, 1991, J Assoc Physicians India	-NA-	1991	29	50	58
Balgir RS, 1996, J Assoc Phys India	-NA-	1996	25	82	30.49
Colah RB, 2015, Indian J Med Res	-NA-	2015	24	34	70.59
Colah R, 2014, Curr Opin Hematol	10.1097/MOH.0000000000000029	2014	20	28	71.43
Patra PK, 2011, J Community Genet	10.1007/s12687-011-0050-4	2011	19	28	67.86
Jain DL, 2012, Hemoglobin	10.3109/03630269.2012.691434	2012	17	26	65.38
Italia Y, 2015, J Med Screen	10.1177/0969141314557372	2015	16	22	72.73
Hockham C, 2018, Sci Rep	10.1038/s41598-018-36077-w	2018	15	18	83.33

Table 4: Publication Output of Journals.

Sl. No.	Journal	No. of Publications	Total Citations	*h*-index
1.	Indian Journal of Haematology and Blood Transfusion	40	179	7
2.	Indian Journal of Paediatrics	39	363	12
3.	Indian Journal of Medical Research	38	456	13
4.	Indian Paediatrics	36	258	8
5.	Journal of Clinical and Diagnostic Research	32	103	7
6.	Haemoglobin	29	352	12
7.	Indian Journal of Pathology and Microbiology	23	122	6
8.	Indian Journal of Clinical Biochemistry	20	97	6
9.	Journal of Association of Physicians of India	20	197	7
10.	Annals of Haematology	15	164	6

Most Productive and Highly Cited Journals

In all, 677 journals are counted to publish articles on "Sickle Cell Anemia" research by Indian researchers. The dataset for most productive journal clearly points the difference between a productive journal and an impactful one. One journal may be productive but simultaneously it can't be impactful. The dataset for total citations and *h*-index indicates uneven or abnormal behavior with respect to number of publications. Of the 10 most productive journals, 3 of them are having *h*-index between 10 to 15 followed 12 journals with average *h*-index between 5 to 10 and 5 journals are having *h*-index between 1 to 5 range [Table 4].

The bibliographic coupling between different journals is depicted in Figure 5 generated with VOSviewer. Multi-colored circles imply the variations in bibliographic coupling. The size of the circles is related proportionally with the

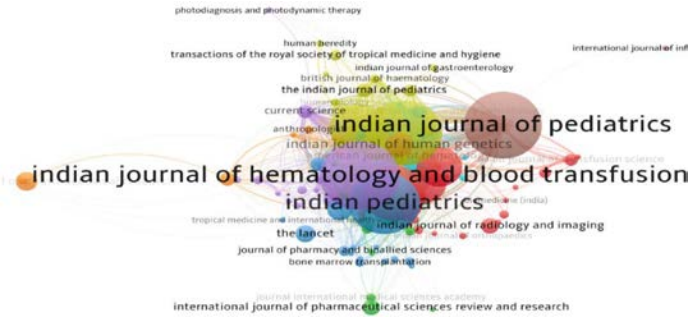

Figure 5: Network map of Bibliographic Coupling of Journals.Highly Cited Works

productivity of the journals. The font size of the letters in the name of the journals is also proportionally related to the number of documents in the journals. The linked lines of similar color depict bibliographic coupling links between the items. Taking the minimum number of documents of a source as 3, the network map is created for the total 698 sources out of which only 111 sources have minimum number of 3 documents which are bibliographically coupled. Out of 111 sources, 106 sources are connected. These 106 items are divided into 8 clusters. Cluster 1 (red) has 24 sources like *Asian Journal of Transfusion Science, Biomedicine (India), Indian Journal of Orthopedics*, *Indian Journal of Pathology and Microbiology* etc. The second cluster (green) has 16 sources, some of them are *International Journal of Pharmaceutical Sciences Review and Research*, *Journal International Medical Sciences Academy*, *Gene, Annals of Hematology* etc. Cluster 3 (deep blue) has 15 journals like The Lancet, Bone Marrow Transplantation, Journal of Pharmacy and Bio allied Sciences, Indian Pediatrics. Cluster 4 (yellow) has 15 sources, prominent are *Hemoglobin, Indian Journal of Medical Research, Human Heredity, British Journal of Hematology*. Cluster 5 (violet) is with 14 sources including *Indian Journal of Hematology and Blood Transfusion, Current Science, PLoS One, Journal of the Indian Medical Association.* Cluster 6 (shallow blue) has 12 sources, four of them are *Current Opinion in Hematology, Scientific Reports, Free Radical*

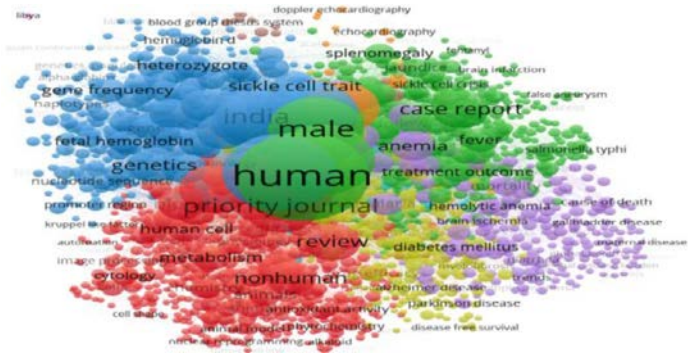

Figure 6: Network of Keyword Co-occurrence.

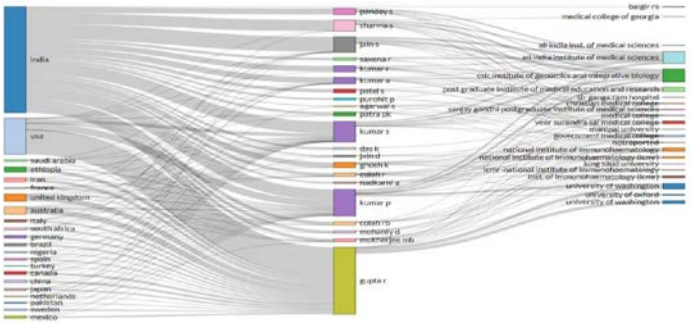

Figure 7: Three field plot showing countries, authors and affiliations.

Research, Prenatal Diagnosis, Pediatric Blood and Cancer. Cluster 7 (orange) has 8 sources, some are *International Journal for Pharmaceutical Research*, *Anthropologist, Online Journal of Health and Allied Sciences, Medical Journal Armed Forces India.* Lastly, the cluster 8 (brown) has only two journals namely, Indian Journal of Pediatrics and International Journal of Infectious Diseases. The number of links between different journals in the network map indicates the strength of the bibliographic coupling between them.

Year-wise distribution of Publication

In all, bibliographic details of 1505 publications on "Sickle Cell Anemia" have been extracted from Scopus and filtered too on the basis of year of publication of those. It has been observed that the distribution of data obtained (after tabulation in ascending order of the year of publication) is not normal. The value of R^2 (≈1) indicates that the chronological growth in the number of publications is consistent [Figure 8]. The descriptive statistical analysis of the annual scientific production indicates the dataset to be skewed and leptokurtic. The indexing information on Scopus initiates from 1958 and it is considered till 2020. From 1958 to 1996 the number of publications counted is in the range of 1–10 except for 1994 it exceeds to 11. Highest in this range is in 1992 and 1995 amounting to 10 publications each. From 1997 till 2015 the number of publications is double digit i.e., in the range 10 to 50, except for 1998 a single digit number 8 is encountered.

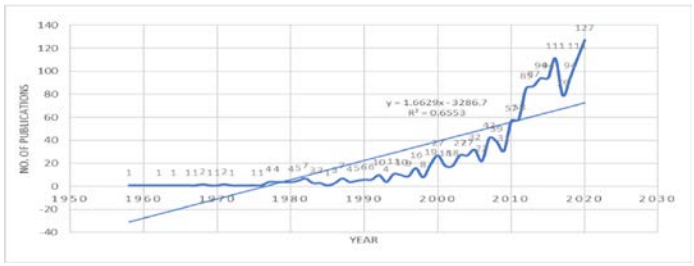

Figure 8: Annual publication output.

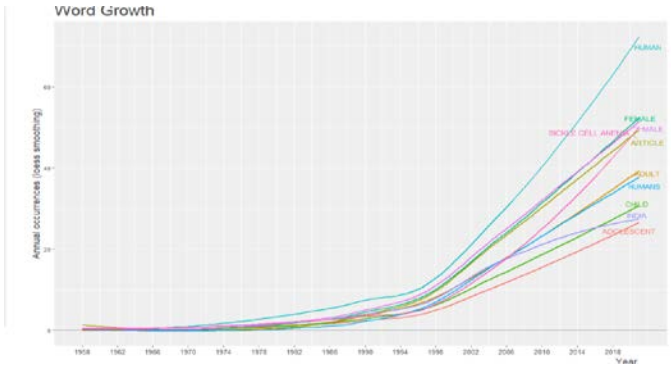

Figure 9: Mapping word-growth of prominent keywords year-wise.

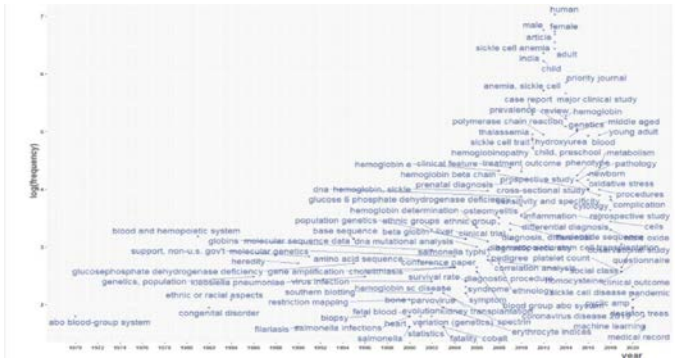

Figure 10: Mapping the trending topics on Sickle-Cell Anemia.

The highest in this range is 42 publications in 2007. From the year 2010 to 2020 there is gradual increase in the number of publications though it is uneven. The number of publications is in the range 50 to 130 (the upper limit of this range is taken as 130 because the overall highest number of publications are observed to be 127 so it is rounded off to 130). The highest number of publications among all years are observed in the year 2020 which is 127 followed by a tie of 111 publications in the year 2016 and 2019.

Significant Keywords and Trending Topics

The keyword co-occurrence in research acts as a secondary support to get an insight into main topics in any subject. Figure 9 and Figure 10 are generated in Biblioshiny web interface through R-Studio platform and Figure 6 is created using the common network visualization tool VOSviewer.

Figure 6 gives a network map of keyword co-occurrences at a glance. The radius of the circle and font size is proportional to the frequency of occurrence of a keyword. The links between the keywords that can be identified with difference in color. Total 12834 keywords were identified, taking minimum number of occurrences of a keyword as 4, it was found that 1937 keywords meet this threshold. These connected keywords are divided into 10 clusters, each with definite number of keywords presented with different colors. Cluster 1 (red) has 580 keywords. Some are *human cell, review, metabolism, cytology, non-human, antioxidant activity, animals*. Cluster 2 (green) has 445 keywords; like *human, male, case report, sickle cell crisis, sickle cell anemia, hemoglobin, differential diagnosis, diagnostic imaging, treatment outcome* etc. Cluster 3 (deep blue) has 323 keywords. Some are *sickle cell trait, fetal hemoglobin, genetics, India, nucleotide sequence, gene frequency, hemoglobin d, heterozygote, alpha globins, haplotypes, gene, gene cluster, genetic linkage, indigenous people, high risk population, tribal people, Indian* etc. Cluster 4 (yellow) has 268 keywords includes *disease free survival, diagnostic value, drug efficacy, hematologic diseases, hematology, survival rate, thalassemia, stem cell mobilization, transplantation, hemophilia, hemosiderosis* etc. Cluster 5 (violet) has 197 keywords in total. Some prominent are *Alzheimer disease, cause of death, cerebral palsy, cardiovascular disease, endocrine disease, gallbladder disease, hepatitis, leukemia, life expectancy, metabolic disorder, newborn disease, pancreas cancer, Parkinson disease, rabies, rheumatoid arthritis, skin disease, stomach cancer, tuberculosis, turner syndrome, urinary tract infection, visceral leishmaniasis, wounds and injuries.* Cluster 6 (shallow blue) has 75 keywords on total. Some of them are *adolescent, sickling, cross sectional study, disease course, gender, hepatitis b virus, heart, immunoglobulin g, pilot project, questionnaire, sample size, virology, virus DNA, statistical analysis, sequence analysis.* Cluster 7 (orange) has 24 keywords in club. Some are *echocardiography, doppler echocardiography, hematologic tests, blood analysis, blood flow velocity, cell count, case control study, eosinophil count, saliva, blood examination, acute albumin level, serum.* Cluster 8 (brown) has 23 keywords in sum *like blood group antigen, blood group rhesus system, cross-sectional study, immunology, blood group typing, blood transfusion reaction, Coombs's test, isoantibodies, tertiary care centers*. Cluster 9 (purple) and Cluster 10 (shallow red) has one keyword each and these *Libya* and *Edetic acid* respectively.

Figure 9, created using the Biblioshiny software with specific graphical parameters: Field as "keyword plus", occurrences "per year", with no confidence interval and the top 10 keywords that are with the maximum frequencies are considered. The keyword human (dark blue) has maximum frequency of occurrence of 1107 followed by male (violet), female (green), article (light brown), Sickle Cell (purple), adult (brown), humans (light blue), India (navy blue), child (dark green), adolescent (red) with frequencies 850, 830,783, 681, 613, 589, 573, 492, 414 respectively, during 1958 to 1990 the keywords have frequency of occurrences as 82, 46, 41, 26, 46, 25, 13, 46, 25, 27 respectively. Again, the keyword human has the highest frequency of occurrence. In 1991 to 2010, human has the maximum frequency of occurrence (366) followed by the word male (309). During 2011 to 2020, similar phenomenon is seen with human having the maximum frequency of occurrence (659) but this time the keyword female follows it with 500 as the frequency of occurrence [Table 5].

Table 5: Frequency of top 10 most occurring keywords in different time slots.

Years	Human	Male	Female	Article	Sickle Cell Anemia	Adult	Humans	India	Child	Adolescent
1958-1990	82	46	41	26	46	25	13	46	25	27
1991-2010	366	309	289	281	196	226	215	215	177	140
2011-2020	659	495	500	476	439	362	361	312	290	247
Total	1107	850	830	783	681	613	589	573	492	414

Figure 10 depicts the occurrences of keywords in publications in different years. The keywords which are most occurring are in the top and which are least occurring are in the end. During 1970 to 1990 keywords like *ABO-Blood Group System, Blood and Hemopoietic system, Congenital Disorder, Genetics, Heredity* were observed with frequencies 6, 24, 7, 12, 15 respectively. From 1991 to 2000 some of the significant keywords are *base sequence, molecular sequence data, restriction mapping, fetal blood, heart, evolution, bone, biopsy* occurring 23, 23, 8, 6, 5, 6, 8 and 6 times respectively. From 2001 to 2010, the mostly occurring keywords are *ethnic groups* (27 occurrences), *globins* (19), *population genetics* (23), *beta globin* (23), *hemoglobin determination* (28), *DNA mutational analysis* (20), *hemoglobin-sickle* (53), *Pedigree* (19), *hemoglobin E* (75), *Osteomyelitis* (33), *clinical feature* (79), *prenatal diagnosis* (60), *treatment outcome* (74), *DNA* (48), *differential diagnosis* (57), *clinical trial* (27). During the decade 2011 to 2020, the prominent keywords are *case report* (198 occurrences), *thalassemia* (141), *sickle cell trait* (131), *hemoglobinopathy* (111), *male* (862), *child* (501), *anemia-sickle cell* (303), *child-preschool* (128), *polymerase chain reaction* (140), *human* (1129), *female* (840), *sickle cell anemia* (695), *adult* (622), *hemoglobin* (195), *genetics* (154), *blood* (137), *young adult* (139), *hemoglobin beta chain* (54), *metabolism* (79), *cytology* (45), *hematopoietic stem cell transplantation* (21), *homocysteine* (13), *platelet count* (13), *machine learning* (6), *medial record* (5). The keywords listed in Table 5 are on the basis of decreasing order frequency of occurrences in the literature.

DISCUSSION

Bibliometric analysis is an effective and efficient tool for knowing the current status and prediction of future

development trends in the knowledge domain of area studied and this makes it different from systematic reviews.[22-23] Research on "Sickle Cell Anemia" comprised of a total of 1505 publications and these are contributed by 320 authors from 98 organizations and collaborated with 120 countries. The average research productivity in the subject was 4.23 authors per document and average citations per year per document is 2.29. The research output (677 productive journals) received total citations of 26543 since publication. The average performance in the subject was 17.59 citations per paper. The USA topped in the global rank to collaborate with India (32.63%) followed by United Kingdom and Australia (6.22% and 5.44%) and the rest 11 most collaborative and productive countries. Moreover, the study provides an insight to key countries, key organizations, key authors, prominent source journals, significant keywords and trending topics of research on "Sickle Cell Anemia". The findings of the study are very significant as it gives a pen picture of research conducted by Indian Scientists and medicos on Sickle Cell Anemia, a deadly genetic disorder of blood. The institutions like ICMR-National Institute of Immunohematology, ICMR-National Institute of Research in Tribal Health and AIIMS are contributing research works prominently in the field. India, being a developing nation has to draw plans to mitigate this genetic disorder through more R&D activities though these activities are increasing in last decades.

Indian Journal of Hematology and Blood Transfusion, Indian Journal of Pediatrics, *Indian Journal of Medical Research* are the most preferred journals as implied by their maximum number of publications. *The Lancet* is found to the most impactful journal with highest number of citations which reveals it's quality. These journals are of high impact, and the quality and publication of quality papers have also raised the academic impact of these journals.

The paper also examines co-authorship in terms of countries, authors and organizations. Co-authorship implies evaluating the relationships among the items (countries, organizations) through the number of co-authored documents. It is applied to assess the cooperation between different organizations, countries and authors in the field of Sickle cell anaemia research. The quantitative assessment of closeness of cooperation is given by indicators links, number of documents, total link strength (TLS).[21] Higher TLS implies that the countries, organizations and authors tend to work more collaboratively than ones with low TLS. That means the greater the value, the more frequent the cooperation. The analysis of cluster also indicates the closeness/cooperation between the countries, organizations and authors. Presence in a similar cluster with common colour indicates the cooperation in research. The findings for this analysis are mentioned in the previous sections.

In the bibliometric study, as mentioned, the frequency of appearance of keywords in a data set reveals the hot spot categories and future development of a discipline.[21] Analysis is performed in three different ways. According to the keyword co-occurrence network and cluster analysis performed by VOSviewer, all identified keywords were extracted from the Scopus database and analysis is done. Dividing the keywords into 10 different clusters on the basis of their commonness in occurrence and relationship among them the cluster and network analysis is performed. Parallelly, using the Biblioshiny, the frequency of occurrences of the keywords is evaluated and most frequently occurring keywords are depicted. Moreover, the trends of the keywords are also analysed. In the last decade, from 2011 to 2020, *thalassemia*, *sickle cell trait*, *hemoglobinopathy*, *male*, *child*, *anemia-sickle cell*, *child-preschool*, *polymerase chain reaction*, *human*, *female*, *sickle cell anemia*, *adult*, *hemoglobin*, *genetics*, *blood*, *young adult*, *hemoglobin beta chain*, *metabolism*, *cytology*, *hematopoietic stem cell transplantation* were the keywords with maximum frequencies implies that focus of the research was human beings with special attention to gender for which keywords like male, female and age group for which child-preschool, young adult came into existence. The research also turned on to state-of-the-art medical treatments like stem cell transplantation. Research in this area has left behind ample scope for new researcher to explore diagnosis and treatment methods to overcome this genetic ailment. The study has been conducted with enriched bibliometric techniques like visualization which is drawing new insights. Moreover, analysis of performance at article, author, institution and country level using citation metrics and science mapping with bibliographic coupling and co-authorship spreads light in this study. Since the study is limited to only network and cluster analysis, so for bibliometricians this study leaves possibilities to conduct more extensive bibliometric analysis using performance-citation related metrics like collaborative coefficient, degree of collaboration, modified collaborative coefficient, authorship pattens to understand the sociological aspects of research in this field which is the future aspect of research in this field. Moreover, the study considers data from only one database i.e., Scopus that too limited to the Indian publications only, leaving opportunity to conduct comparative study with data from multiple databases considering the global publications.

CONCLUSION

This paper contributes in three ways in expanding information on research related to Sickle Cell Anemia. First, it gives insights to previous work published on the topic which helps to find the research gap. Second, it conducts a bibliometric and network analysis to discover the most impactful articles, co-authorship of countries, organizations and authors, highly cited publications, preferred journals, keyword analysis and

chronological distribution of publications. Third, analysis of keywords depicts the trending topics and give meaning directions for future research. This information generated in this study will be of immense help for researchers working in this area.

In a nutshell, it can be mentioned that collaboration of India is though more with nations like the USA, U.K. and Australia in terms of number of publications but taking the angle of occurrence of co-citation India is much mutual to nations like Canada, Jamaica though the USA is also in between these two nations. In a nutshell, the strength of international collaboration in "Sickle Cell Anemia" research is observed to be highest among the USA, U.K. and India vis-à-vis other top 26 countries. India is developing in every aspect. Be it medical science, space science etc. but findings of this study still depict India to be remaining behind the global scenario on research in "Sickle Cell Anemia". As per a survey by ICMR, 20% of children with SCD are killed at the age of two and 30% of children with SCD dies before attaining adulthood. This is a very terrible sign for future of India. This is reflected in the findings of the study also. The collaboration network of India is not too strong than that of the USA, moreover India also lags behind in number of publications. Authors like Colah B. Roshan and Ghosh Kanjakshsa of ICMR- National Institute of Immunohematology are contributing quality research work in global standards but these are needed to be practically implemented in the field. This could be achieved only through implementation of government initiatives and policies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Wade C. Impact of sickle cell disease on executive functioning in an adult population. Dissertation Abstracts International: Section B: The Sciences and Engineering. 2020;81:(10-B).
2. Pecker LH, Lanzkron S. Sickle cell disease. Annals of internal medicine. 2021;174(1):ITC1-16. doi: 10.7326/AITC202101190, PMID 33428443.
3. Gardner RV. Sickle cell disease: Advances in treatment. The Ochsner Journal. 2018;18(4):377-89. doi: 10.31486/toj.18.0076, PMID 30559624.
4. Jain D, Italia K, Sarathi V, Ghosh K, Colah R. Sickle cell anemia from central India: A retrospective analysis. Indian Pediatrics. 2012;49(11):911-3. doi: 10.1007/s13312-012-0217-z, PMID 22728629.
5. Sheshadri V, Shabeer P, Santhirapala V, Jayaram A, Krishnamurti L, Menon N. Mortality in sickle cell disease: A population-based study in an aboriginal community in the Gudalur Valley, Nilgiris, Tamil Nadu, India. Pediatric Blood and Cancer. 2021;68(3):e28875. doi: 10.1002/pbc.28875, PMID 33381914.
6. Naheem KT, Nagalingam U, Ramesha B. Chronic liver disease (CLD) research in SAARC countries: A scientometric analysis of research output during 1996-2015. Annals of Library and Information Studies. 2017;64(1):59-68.
7. Ekundayo TC, Okoh AI. A global bibliometric analysis of Plesiomonas-related research (1990-2017). PLOS ONE. 2018;13(11):e0207655. doi: 10.1371/journal.pone.0207655, PMID 30496198.
8. Yang W, Zhang J, Ma R. The prediction of infectious diseases: A bibliometric analysis. International Journal of Environmental Research and Public Health. 2020;17(17):6218. doi: 10.3390/ijerph17176218, PMID 32867133.
9. Ram S. Coronavirus research trends: A 50–year bibliometric assessment. Science and Technology Libraries. 2020;39(2):210-26. doi: 10.1080/0194262X.2020.1742270.
10. Gupta BM, Dhawan SM, Ahmed KM, Mamdapur GM. Use of Convalescent Plasma Therapy for COVID-19: A Scientometric Assessment of Global Publications during 2020-21.
11. Adesina OA, Opesade AO. Bibliometric Analysis of Sickle Cell anemia Literature on Nigeria listed in PubMed between 2006 and 2016. Library Philosophy and Practice. 2018;1923.
12. Okoroiwu HU, López-Muñoz F, Povedano-Montero FJ. Bibliometric analysis of global sickle cell disease research from 1997 to 2017. Hematology, Transfusion and Cell Therapy. 2020:S2513-1379(20):31305-5. doi: 10.1016/j.htct.2020.09.156, PMID 33423980.
13. Musa HH, El-Sharief ME, Musa IH, Musa TH, Akintunde TY. Global Scientific Research Output on sickle cell disease: A comprehensive bibliometric analysis of web of Science Publications. Scientific African. 2021;12:(e00774). doi: 10.1016/j.sciaf.2021.e00774.
14. Chen C, Dubin R, Kim MC. Emerging trends and new developments in regenerative medicine: A scientometric update (2000 - 2014). Expert Opinion on Biological Therapy. 2014;14(9):1295-317. doi: 10.1517/14712598.2014.920813, PMID 25077605.
15. Hong Y, Yao Q, Yang Y, Feng JJ, Wu SD, Ji WX, *et al*. Knowledge structure and theme trends analysis on general practitioner research: A co-word perspective. BMC Family Practice. 2016;17(10):10. doi: 10.1186/s12875-016-0403-5, PMID 26831329.
16. van ECK NJ, Waltman L. Software survey: VOSviewer, a Computer Program for Bibliometric Mapping. Scientometrics. 2010;84(2):523-38. doi: 10.1007/s11192-009-0146-3, PMID 20585380.
17. Aria M, Cuccurullo CC. bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics. 2017;11(4):959-75. doi: 10.1016/j.joi.2017.08.007.
18. Synnestvedt MB, Chen C, Holmes JH. CiteSpace II: Visualization and knowledge discovery in bibliographic databases. AMIA Annual Symposium Proceedings. 2005:724-28. PMID 16779135.
19. Gu D, Li T, Wang X, Yang X, Yu Z. Visualizing the intellectual structure and evolution of electronic health and telemedicine research. International Journal of Medical Informatics. 2019;130:103947. doi: 10.1016/j.ijmedinf.2019.08.007.
20. Laengle S, Merigó JM, Modak NM, Yang JB. Bibliometrics in operations research and management science: A university analysis. Annals of Operation Research. 2020;294(1-2):769-813. doi: 10.1007/s10479-018-3017-6.
21. Wu H, Tong L, Wang Y, Yan H, Sun Z. Bibliometric Analysis of Global Research Trends on Ultrasound Microbubble: A Quickly Developing Field. Frontiers in Pharmacology. 2021;12:646626. doi: 10.3389/fphar.2021.646626. PMID 33967783.
22. Grant MJ, Booth A. A typology of reviews: An analysis of 14 review types and associated methodologies. Health Information and Libraries Journal. 2009;26(2):91-108. doi: 10.1111/j.1471-1842.2009.00848.x, PMID 19490148.
23. Møller AM, Myles PS. What makes a good systematic review and meta-analysis? British Journal of Anaesthesia. 2016;117(4):428-30. doi: 10.1093/bja/aew264, PMID 28077528.