BIBLIOMETRIC ANALYSIS OF GLOBAL SICKLE CELL DISEASE RESEARCH FROM 1997 – 2017

Henshaw Okoroiwu (okoroiwuhenshaw@gmail.com)
University of Calabar https://orcid.org/0000-0003-2658-3788
Francisco López-Muñoz
University Camelo Jose Cela
F. Javier Povedano-Montero
Universidad Europe de Madrid

Research article

Keywords: Sickle cell, Sickle cell disease, sickle cell anemia, Bibliometric analysis

Posted Date: June 17th, 2019

DOI: https://doi.org/10.21203/rs.2.10361/v1

Version of Record: A version of this preprint was published at Hematology, Transfusion and Cell Therapy on December 1st, 2020. See the published version at https://doi.org/10.1016/j.htct.2020.09.156.

Abstract

Background: Sickle cell disease (SCD) is an autosomal recessive genetic disease caused by single point mutation in the β-globin chain of the hemoglobin. It has been recognized by World Health Organization as a public health priority since 2006. Methods: Scopus database was used in this study with the search descriptors: “sickle cell” and “sickle cell disease”. We applied common bibliometric indicators to evaluate the trend in scientific literature in sickle cell disease research. Results: We retrieved a total of 19,921 scientific literatures in the repertoire from 1997 to 2017. Price law was fulfilled in the trend of production of scientific literature in SCD as the growth of scientific literature was more exponential (r = 0.959) than linear (r = 0.9449). We observed a duplication time of 4.55 years. The Bradford core was made up of 69 journals with Blood at the top, publishing the most number of articles. The most productive institutions were mostly United States agencies and hospitals. United States was the most productive country. National Institute of Health was the most productive institution and also had the highest number of citation. Vichinsky E was the most productive author while the most cited article was published by Circulation. Conclusion: The growth of scientific literature in SCD was found to be high. However, the exponential growth trend shows a “yet-to-be-explored” area of research. This study will be useful for physicians, researchers, research funders, policy cum decision makers.

Background

Sickle cell disease (SCD) was first identified in 1910 by James Herrick [1] followed by description of the molecular basis of the disease in 1949 by Linus Pauling and colleagues [2,3,4]. It is an autosomal recessive genetic disease cause by a single point mutation (69 A>T) in the β-globin chain of the hemoglobin in which there is substitution of hydrophobic valine residue for negatively charged glutamic acid residue, consequently resulting in non-covalent polymerization and double strand formation under low oxygen [3,5]. This non-covalent polymerization of the sickle hemoglobin into long fibres under oxygen tension (deoxygenation) is the major pathological event in SCD. The resulting fibres distort red blood cells into atypical and heterogeneous shapes; crescent (classical sickle), elongated, granular and oval shapes [6,7,8] that lack deformability. These red cells with decreased deformability (loss of membrane elasticity) are rigid and sticky, and are usually trapped in narrow capillary blood vessels causing frequent episodes of vaso-occlusion and ischemia [6,9]. More so, the stiff cells that are unable to return to the normal shape are spotted and destroyed via hemolysis [9,10]. SCD encompasses sickle cell anemia (which is the inheritance of two S genes), and co-inheritance of hemoglobin C (HbSC) or β-thalassaemia (HbS/β-thalassaemia). However, sickle cell anemia is the most prevalent globally, accounting for an estimated 83% of all newborn with SCD [11]. Conversely, those who inherit single HBS gene together with HbA (heterozygotes) are carriers and usually asymptomatic.

The World Health Organization in 2006 recognized sickle cell anemia as a public health priority and in 63rd (2010) World Health Assembly adopted resolution on the prevention and management of birth defects including sickle cell disease and thalassaemia [12,13,14]. Sickle cell disease prevalence depends on the sickle cell trait. The sickle cell trait is now widespread, reaching its highest prevalence in parts of

Africa, the Mediterranean basin and Saudi Arabia. In countries such as Ghana, Cameroon, Nigeria, Gabon and Republic of Congo, the prevalence is between 20% and 30% while in some parts of Uganda, it is up to 45%. In countries where the trait is above 20%, the disease affects about 2% of the population [15]. In a global map of HbS allele frequency distribution with Bayesian geostatistical model using a database of sickle hemoglobin surveys by Piel and colleagues, [16] it was shown that 50% of total AS and SS neonates were born in only three countries; Nigeria, India and Democratic Republic of Congo. Nigeria has the highest incidence of SCD in the world with approximately 91,011 children born with the defect accounting for approximately 2% of all newborns annually, followed by Democratic Republic of Congo with a 39,743 sickle cell births per year [17].

“Bibliometrics” is a term coined by Allan Pritchad to define the use of mathematical cum statistical procedures to the process of propagation of written communication in field of scientific discipline through the quantitative study of varying aspect of this type of communication [18,19]. Bibliometric studies are relevant tools in social and scientific evaluation of a given discipline within a specified time frame. Bibliometric indicators are proxy markers for activity in a field of research [20,21].

This study was aimed at identifying the trend in sickle cell disease as well as analyzes the structure of the evolving sickle cell disease research community network over time. The result of this study will be relevant to clinicians, researchers, government/health policy makers as well as research finders.

Methods

Data Source

Scopus database was used for this bibliometric study considering that it is the largest abstract and citation database of peer reviewed literature. It index approximately 22,000 journal titles from over 5000 publishers. Among these, approximately 20,000 are peer-reviewed journals across the scientific, medical, technical and social science disciplines. In comparison with other bibliometric database, Scopus in more suited to biomedical field as it is comprehensive, user friendly coupled with the fact that it largely regarded as the world’s largest database for abstract and citation information and is conveniently used in many bibliometric studies [22,23]. We used remote downloading technique to retrieve articles published from 1997 to 2017 containing the descriptors *“sickle cell”, “sickle cell disease”* limited to three fields; title, keyword, and/or abstract. We included all original articles, reviews, editorials, brief reports letters to editors and so forth.

Bibliometric indicators

The bibliometric indicators used in this study is similar to those used in our previous bibliometric studies [19,24]. These includes: Price’s law, duplication time and annual growth rate, Bradford zones, Bradford zones, bibliometric coupling, key word analysis.

We employed Price's law [25] as an indicator of productivity as it is the most common bibliometric indicator for assessing productivity within a specific discipline or country. Price law uses exponential growth evaluation. In order to assess if scientific production in sickle cell disease follows Price's law of exponential growth, we modeled our generated data into linear adjustment utilizing the equation $y = 50.904x - 101216$ and another adjustment to an exponential curve using the equation $y = 3E - 46e^{0.0557x}$. Price law is fulfilled when the coefficient of determination of the exponential curve is greater than that of the linear curve.

Duplication time and annual growth rate was also utilized in this study as bibliometric indicators. Duplication rime refers to the time (in years) it takes to a subject to double its production. Conversely, annual growth rate conveys information on the present growth in relation to the proceeding year and it is denoted in percentage. The duplication time was calculated using the formula:

Formula 1

where b represents the constant that relates the growth rate with the already acquired output of the discipline. The annual growth rate was calculated using the formula:

$R = 100 (e^b - 1)$

Bradford's law [26] was utilized to determine the dispersion of scientific literature. Bradford created concentric zones of productivity referred to as Bradford zones with decreasing density of information. He hypothesized that each zone contains similar number of documents. Whereas the number of journals in which they are produced increases as one moves from one zone to the next. The division of journals in the different Bradford zones is as follows: 1, n, n^2 … The number of articles are divided into 3 groups of approximately the same size where one is the core zone while the other two are peripheral zones. This stratification aids in identifying the most widely used and highest impact journals in a specific area of interest/evaluation.

Impact factor (IF) was used as a measure of the journal's influence. Impact Factor was originally developed by Institute for Science Information (Philadelphia, PA, USA) as a bibliometric indicator and updated annually in the *Journal Citation Report* (JCR) section of Science Citation Index Expanded (SCI). The calculation takes into account the number of times a given journal is cited by SCI journals within the two preceding years. The value is usually a marker of scientific "prestige". We used the impact factor data of 2018 by JCR for this study.

National participation index (PaI) in overall scientific publication in sickle cell disease and in the fields of medicine, and other medical cum health related disciplines in world's ten most productive countries in sickle cell disease research during the period 1997 to 2017 was evaluated. It is the number of documents on the topic in question (in this case sickle cell) compared with global participation index in biomedical and sciences more widely and other subareas and medical and health sciences. We further correlated the participation index with health data such as per capita expenditure on health as well as the country's

gross domestic expenditure on research with data obtained from the World Bank [27] and World Health Organization Department of Health Statistics [28], being data of 2015.

We also used bibliometric coupling to assess the trend of interest of various institutions involved in sickle cell disease research.

Key words analysis was used to evaluate the trend of discussion and research with respect to disease characteristics, clinical research, pathology and treatment and effect.

Results

Assessment of global publication

A total of 19,921 research publications were recovered using the search criteria within the study period (1997 - 2017). Of these, 63.72% (n=12,693) were original articles, while 17.35% (n=3,457), 5.87% (n=1,170), 3.21% (n=639), 2.79% (n=555), 2.41% (n=480), 2.15% (n=429), 1.52% (n=302), 0.47% (n=94), 0.36% (n=72), 0.13% (n=25) and 0.03% (n=5) were Reviews, Letters, Notes, Conference papers, Editorials, Book chapters, Short surveys, Errata, Articles in press, Books and Conference reviews, respectively. (Table 1)

Table 1

As shown in Figure 1, there has been remarkable increase in the number of scientific publications in the area of sickle cell disease research over the 20 years lapse. The linear and exponential adjustments according to figure 1 yielded correlation coefficient *r* of 0.9449 and 0.959, respectively. This shows that the growth of scientific publication in area of SCD is in exponential growth stage (as *r* for exponential curve > *r* for linear curve) with an average of annual increase of 4.93%. It is worthy of note that only 4.1% data is not explained by the equation.

Figure 2 shows the temporal production of scientific literature. In order to calculate the duplication time, the trend line was fitted to the equation $y = 1129.2e^{0.152x}$ with a correlation coefficient 0.9590 over the 29 years time frame. Applying the equation for duplication times gives a duplication time of 4.52 years. This means that production of scientific publications doubles every 4.55 years.

Figure 3 shows that the global production of scientific literature in area of SCD stratified into five years period. It could be observed that in each 5-years period, there is a gradual increase over the previous one. It was observed that the 2012 - 2017 periods contains majority of the documents accounting for 40.5% of the total documents.

Table 2 shows distribution of the journals in Bradford zones. A total of 4,103 different journals published the article under study. However, only 13.67% of them were responsible for more than 60% (66.36%) of the published manuscripts. The core Bradford zone (those containing the largest number of articles)

consisted of 69 journals; notably *Blood*, *American Journal of Hematology*, *British Journal of Hematology*, and others.

Table 2: Bradford distribution of journals

Analysis of sources with highest publication

Table 3 shows analysis of sources with the largest number of publications in SCD research represented by the top 10 journals in the list with their corresponding impact factors according to JCR 2018 and their productivity index in total database within the study period. Seven (7) and 4 of the journals have impact factors greater than 2 and 5 respectively. *Blood*, *American Journal of Hematology*, *British Journal of Hematology*, *Pediatric Blood and Cancer*, *Journal of Pediatric Hematology and Oncology*, *Hemoglobin*, *Revista Brasileira de Hematologia e Hemoterapia*, *Transfusion*, *Haematologia* and *PLoS ONE* were the top 10 most productive journals

Table 3: Analysis of sources with the highest number of publications

Table 4 shows the top 20 most productive countries in SCD research. United States (n = 9167) and United Kingdom (n = 1,869) with participation index (PI) of 46.02 and 9.38, respectively. Next on the list are France (n = 1379; PI = 6.92) Brazil (n = 933; PI = 4.68), India (n = 660; PI = 3.31), Italy (n = 577; PI = 2.90), Nigeria (n = 574; PI = 2.88), and Canada (n = 522; PI = 2.62). It is pertinent to note the only United State accounted for nearly half (46.02%) of all publications.

Table 4: Top 20 most productive countries.

Comparative analysis of the productivity of the top 10 most productive countries in sickle cell research with their overall production in the area of medicine showed that only 7 out of the 10 countries (United Kingdom, USA, France, Brazil, India, Nigeria and Saudi Arabia) devoted a higher proportion of attention to the study of SCD in relation to medicine in general. United States, United Kingdom and France were consistently leading both in sickle cell disease research as well as research in medicine generally (Table 5).

Table 5: Comparative analysis of the productivity of the top 10 most productive countries in SCD research with their overall production in the area of medicine and other areas of biomedical research.

Figure 4 shows the correlation between production of scientific literature on SCD and current health expenditure (CHE) and gross domestic expenditure on research in 15 most productive countries in SCD research. In the analysis of the correlation between the participation index and the current health expenditure of each of the top 15 most productive countries in SCD, the trend obtained was inconsistent, except United States and France that maintained their trend. More so, the correlation between 15 top producers and gross domestic expenditure on research also showed inconsistent pattern except for France.

Productivity of Institutions

The top most productive institutions in SCD research were represented in Table 8. National Institute of Health, Bethesda was the most productive institution (n = 447; 2.24%) followed by INSERM (n = 396; 1.99%), Children Hospital of Philadelphia (n = 350; 1.76%), Harvard Medical School (n = 310; 1.56%), St. Jude's Children Research Hospital (n = 306; 1.54%) and others.

Seventy percent (70%) of the top 20 most productive institutions are resident in the United States, whereas the rest are from France (15%), and UK (15%). Five (25%) of the topmost productive institutions are public universities, while 5 (25%), 4 (20%), 3 (15%), 2 (10%) and 1 (5%) are private universities/medical colleges, not-for-profit hospitals, government agencies, public hospital and private hospital, respectively (Table 6).

Table 6: Top 20 most productive institutions

Citation analysis of top most productive institutions showed the highest citation in National Institute of Health, Bethesda and National Health, Lung and Blood Institute while the highest citation per document was recorded in National Heart, Lung and Blood Institute and Duke University School of Medicine (Table 7).

Table 7: Citation analysis of the 10 top most productive institutions.

Figure 5 shows the bibliometric coupling between most productive institutions in SCD research.

Productivity of Authors

Table 8 shows analysis of top 10 most productive authors in SCD research. Vichinsky E, DeBaun MR, Gladwin MT, Steinberg MH, Ballas Sk, Galacteros F, Ware RE, Kato GJ, Wang WC and Connes P were the top 10 productive authors. Eight (8) out of 10 of the most productive authors are resident in the United States and are affiliated to UCSF Benioff Children Hospital, Washington School of Medicine, Pittsburg Heart, Lung, Blood and Vascular Medicine Institute, Thomas Jefferson University, Cincinnati Children Medical Centre and St. Jude's Children Research Hospital. The other two authors are affiliated with Henri-Mondor Hospital and Institute Universitaire de France/Universite Claude Bernard Lyon 1 Villeubanne/Laboritoire d'Excellence du Globule Rouge, all in France.

Table 8: Ten most productive authors in SCD research

Citation of Analysis Articles

Table 9 shows top 10 most cited articles. The most cited articles was an update report dedicated on heart disease and stroke while the second was dedicated to causes of mortality (SCD inclusive). The 3rd top

cited article was dedicated to glutathione metabolism with mention of how oxidative stress plays critical role in pathogenesis of sickle cell disease. Only the 8th top article was dedicated to treatment of SCD using mouse model.

Table 9: Top 10 most cited articles in SCD research

Keyword Analysis

We identified 160 keyword in the search repertoire. Stratification of the keywords into 3 clusters ("*characteristics*", "*clinical research*", "*pathogenesis*", and "*Treatment and effect*") showed that "*characteristics*" composed 24.3%, while "*clinical research*", "*pathogenesis*" and "*treatment and effect*" composed 26.5%, 40.0% and 9.4%, respectively. Under "*characteristics*", "Human" was the most commonly used keyword while "hydroxyurea" and "blood transfusion" were the commonly used keywords under "*treatment and effect*". Controlled study was the most commonly used keyword under "*clinical research*" while anemia was the most commonly used keyword under "*pathogenesis*" (Figure 6).

Discussion

The most utilized document type used in this study by authors is original article which accounted for about 63.72% of all article type used. This is a reflection that the subject matter is experimental or clinical research.

We observed an exponential growth trend in publications relating to SCD research in the past 20 years with an average annual increase of 4.93%, without evidence of reaching saturation as postulated by Price's law, hence fulfilling Price's law. This significant growth in the field of SCD research takes us to the conclusion that the field of SCD is still at the prime of development from the clinical and basic research perspective. This immense growth in research output can be attributed to the continual interest in research geared towards providing curative and disease modifying agents for SCD. There have been immense attempt in diverse clinical trials on stem cell replacement/modification, antisickling agents, gene therapy, antioxidant therapy and other anti adhesive therapies [29]. The observed trend is similar to previous reports in areas of optometry, bipolar disorder and obstetrics and gynecology [24,30,31]. However, the trend is at variance with a linear trend observed in a study in infectious disease (Lassa fever) [19].

Bradford stratification of the articles showed that only 19 journals (1.68%) were responsible for the production of 33.49% of the published literature. This trend reflects a high concentration of publications by a small group of journals. Individual analysis showed that *Blood* had the highest number of publications accounting for approximately 3.22% of all publications followed by *American Journal of Hematology* (2.64%) and *British Journal of Hematology*. All the top ten journals were all dedicated to hematology except *PLoS ONE* that is a multi-disciplinary journal.

Country-wise analysis of research output in sickle cell disease showed that United States, United Kingdom and France topped the research output in SCD research and accounted for 62.32% of total research output in sickle cell disease. Only the United State accounted for 46.02% of the total research output. The reason for this increased productivity cannot be far-fetched as US, UK and France housed the most productive institutions; research institutes, hospitals and universities. More so, the trio are home to major pharmaceutical companies that manufacture the disease modifying drug hydroxyurea approved for the management of SCD: *Droxia*®,*Hydrea*®, and *Litalir*® by Bristol-Myers Squibb Co, USA, France, Italy, India, China, England, Belgium; *Siklos*®, by Addmedica France. India also, is home to the brands of hydroxyurea: *Ribore* (Khandelwal Laboratory Pvt Ltd), *Readrox* (Miracalus Pharma Pvt Ltd), *Neodrea* (Vhb Life Science Inc.), and *Myelostat* (Zydus Cardila), while Italy in addition to Bristol-Myers Squibb Co, housed Teofarma Srl that manufacture *Oncocarbide*™, a brand of hydroxyrea. More so, United States is the home of Emmaus Medical Inc., the pharmaceutical company that manufactured the newly approved disease modifying drug, *Endari*, for SCD. It is pertinent to note that Nigeria was the only African country present in the top 20 most productive countries. It is documented that Nigeria has the highest incidence of sickle cell disease in the world. Hence, the endemicity of disease could be a boost in interest in the study of the disease. More so, Nigeria houses the pharmaceutical company Bond Chemical Pvt Ltd, that manufacture the brand of hydroxyurea, *Oxyurea*. While comparing the most productive countries in SCD research in relation to general medical research and some other infectious disease, the striking interest of Brazil, France, Nigeria and Saudi Arabia were remarkable. More so, the low comparative participation index of Germany in SCD research in relation to general research in medicine was also remarkable.

A comparative analysis between production in SCD and current health expenditure showed that the more the current health expenditure, the greater the productivity in research in SCD with some exceptions: France, Italy, Canada, Germany, Saudi Arabia, Netherlands, Turkey, Jamaica, Greece and Spain where the current health expenditure exceeded their participation index in SDC research. However, Nigeria had the lowest current health expenditure (3.57% of GDP).

Correlation analysis of between scientific productivity in SCD and gross domestic expenditure on research and development (R&D) placed Jamaica, Nigeria and India in the last three positions in the evaluated countries. It is pertinent to note that only 2 of the most productive countries in SCD research (Germany; 2.88% and US; 2.79%) had a gross expenditure on R&D above the Organization for Economic Co-operation and Development (OECD) average of 2.3% while only 4 (Germany; 2.88%, US; 2.79%, France; 2.23% and Netherlands) had a gross expenditure on R&D above the EU28 average of 1.9%. A country's scientific output is a reflection of its earlier investment in research and development in preceding years before the year of analysis and not as a result of a particular economic event in the evaluation period [30,32].

Institution in the United State dominated the list of top 20 most productive institution followed by France and United Kingdom. This further explains why the United States topped in the productivity of SCD research. This finding implies that creation of first class research institutes is fundamental in improving research and academic output of a country. National Institute of Health is a part of the United States

Department of Health and Human Services that serves as the nation's medical agency. It conducts its own research via its Intramural Research Program as well as provides major biomedical research funding to external research facilities via its Extramural Research Program [33]. On the other hand Institut National de la saite et la Recherche Medicale (INSERM) is the French national institute of health and medical research dedicated to human health with the objective to promote health by advancing knowledge about life and disease, treatment innovation public health research [34].

Unlike co-citation analysis which is a measure of the frequency with which two documents are cited together by one document [35], bibliometric coupling occurs when two different studies reference a common third study in their bibliography [36,37]. Bibliographic coupling shows similarity of subject matter of the two works. In this study, we observed the existence of several groups of institutions bibliographically coupled forming a huge bibliographic network. For instance, National Heart, Lung and Blood Institute, St. Jude Children Research Hospital, Children Hospital Boston and John Hopkins School of Medicine can be tagged a premium network of American Institutions. Another group of American institutions that are bibliographically coupled are Harvard Medical School, Medical College of Wisconsin and Cincinnati Children's Hospital Medical Centre. Among the European institutions, INSERM, University College London, were bibliometrically coupled. These institutions citing papers in common reflects a high bibliometric network showing research focused on related topics.

Vichinsky E, DeBaun MR, Gladwin MT, Steinberg MH, Ballas Sk, Galacteros F, Ware RE, Kato GJ, Wang WC and Connes P were the top 10 productive authors. Vichinsky E. focused on treatment, pathogenesis and population based survey especially on chronic organ failure in sickle cell disease [38,39]. DeBaun MR focused on pathogenesis and treatment and affect in SCD [40,41] while Gladwin MT, Steinberg MH, Ballas SK and Galacteros F dedicated most of their studies in treatment/outcome and pathophysiology of SCD. These authors have published several novel works and are likely to have tremendous impact in future development in SCD research.

The core journals utilized in the publication of SCD research were *Blood*, *American Journal of Hematology*, *British Journal of Hematology*, *Pediatric Blood and Cancer*, *Journal of Pediatric Hematology/Oncology*, *Hemoglobin*, *Revista Brazileira de Hematologia e Homoterapia*, *Transfusion*, *Haematologia* and *PLoS ONE*. Subsequently, researchers may pay attention to these journals as future breakthrough in SCD are more likely to be published in these journals.

The article "Heart disease and stroke statistics-2012 updates" a report from American heart association published by *Circulation* was the most cited article accounting for 1% of the total citation. It was an article dedicated to statistics, risk factor and economic costs of heart diseases and stroke. The second most cited article was an article on glutathione and its implications in varying health conditions including sickle cell. The 8th most cited article; "Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin" [42] published by *Science* was the only article among the top 10 articles dedicated solely to sickle cell disease research. Two of the journals (*New England of Medicine* and

Science) are delayed open access while the rest 8 are hybrid journals. These journals are all well established journals with impact factors greater than 4.

Analysis of keywords is an important bibliometric tool used in monitoring and discovering directions and popular topics in research [43,44]. In the present study, there is a shift in terms of research focus from "*characteristics*" to "*pathogenesis*" and "*clinical research*". The cluster of treatment and effect was still low. This suggests a non-saturated research in treatment options "begging" for more research. The rule of translational medicine is usually movement from disease characteristics to treatment [45]. The only disease modifying therapies approved for SCD are hydroxyurea approved in 1997 and L-glutamate approved two decades after (2017). There have been promising studies on antisickling agents as well as gene therapy. Allogenic hematopoietic stem cell transplant and gene therapy are the only curative treatments and has shown promising results [29]. The key advantage of gene therapy over hematopoietic stem cell transplant is the option to use autologous stem cell, hence, shunting the need to screen donors. In the cluster of "*characteristics*", human was the most commonly used keyword while controlled study was the most commonly used keyword in the cluster of "*clinical research*". Anemia was the most commonly used keyword under "pathogenesis" cluster while hydroxyurea was the most commonly used keyword under "treatment and effect" cluster.

However, there are some limitations to this study which are mainly inherent in bibliometric studies. Scopus database was used for this study, hence, local journals that were not indexed in Scopus during the study period were not included. We might have excluded articles on sickle cell if the authors did not include our search descriptors in the title, abstract or keyword. More so, the criteria mapped out by the database themselves determine the subsequent development of studied material [46].

Conclusion

Despite the inherent limitations outlined we believe that this study have provided a significant representation of research trend in sickle cell disease research globally. This study showed that research in SCD is still at exponential stage. Considering the lag in treatment and effect amidst the exponential growth, it is a pointer that research in SCD will most probably increase in growth in proceeding years bearing in mind that curative approach to SCD is still at clinical trial stage. A good number of the research publications were from high income countries including US, UK and France.

List Of Abbreviations

SCD: Sickle cell disease

USA: United States of America

UK: United Kingdom

PL: Productivity level

JCR: Journal citation report

SCI: Science citation index

PaI: Participation index

IF: Impact factor

Declarations

Acknowledgement: Not applicable.

Ethics approval and consent to participate

This study is based analysis from secondary data, thus, did not require ethical clearance.

Consent to publish

Not applicable.

Availability of data and material

Datasets generated and analysed in this study are within the article. The primary source of data, SCOPUS is publicly available.

Competing interest

The authors declare no conflict of interest

Funding

The authors declare that they did not receive funding for this research from any source.

Acknowledgement

Not applicable

Authors' contributions: HUO conceived the study, analysed data, performed literature search and prepared the manuscript; FLM performed database analysis and data curation, analysed data and edited the initial the manuscript; FJP performed database analysis and data curation, analysed data and edited the initial the manuscript; All authors read and approved the final manuscript.

References

1. Herrick JB. Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. JAMA. 1910; 312:1063.
2. Pauling L, Itano HA, Singer SJ, Wells IC. Sickle cell anemia, a molecular disease. Science. 1949; 543-548.
3. Darrow MC, Zhang Y, Cinquin BP, Smith EA, Boudreau R, Rochat RH. Visualizing red blood cell sickling and the effects of inhibition of sphingosine kinase 1 using soft x-ray tomography. Journal of Cell Science 2016; 129: 3511-3517.
4. Hsieh MA, Tisdale JF. Hematopoietic stem cell mobilization with plerixafor in sickle cell disease. Haematologia. 2016; 103(5): 749 -750.
5. Wen J, Tao W, Hao S, Zu Y. Cellular function reinstitution of offspring red blood cells cloned from the sickle cell disease patient blood post CRISPR genome editing. Journal of Hematology and Oncology. 2017; 10: 119.
6. Li X, Du E, Dao M, Suresh S, Karniadakis GE. Patient-specific modeling of individual sickle cell behaviour under transient hypoxia. PLOS Comput Biol. 13(3): e1005426.
7. Kaul DK, Fabry ME, Windsch P, Baez S, Nagel RL. Erythrocytes in sickle cell anemia are heterogeneous in their rheological and hemodynamic characteristics. J Clin Invest, 1983; 72: 22-31.
8. Christopher GW, Hofrichter J, Eaton WA. Understanding the shape of sickle cells. Biophys J. 2005; 88: 1371-1376.
9. Rees DC, Williams TN, Gladwin MT. Sickle-cell disease. Lancet. 2010; 376: 2018-2031.
10. Frenette PS, Atweh GF. Sickle cell disease: old discoveries, new concepts, and future promise. J Clin Invest. 2007; 117: 850-858.
11. Modell B, Darlison M, Global epidemiology of hemoglobin disorders and derived service indicators. Bull World Health Organ. 2008; 86:480-487.
12. World Health Organization. Fifty-ninth World Health Assembly: resolutions and decisions, annex 2010. WHA59/2006/REC/1. Geneva: World Health Organization.
13. World Health Organization. Sixty-third World Health Assembly: resolutions and decisions annexes. WHA63/2010/REC1. Geneva: World Health Organization.
14. Piel FB Hay SI, Gupta S, Weatherall DJ, Williams TN. Global burden of sickle cell anemia in children under five, 2010-2050: modeling based on demographics, Excess mortality, and interventions. PLOS Med. 2013; 10(7): e1001484.

15. World Health Organization. Regional Committee for Africa; Sickle cell disease in the African Region: Fifty-sixth session. 2006. AFR/RC56/17.
16. Piel FB, Patil AP, Howes RE, Nyangiri OA, Gething PW, Dewi M, et al. Global epidemiology of sickle hemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. Lancet. 2013; 381: 142-151.
17. World Atlas. World facts; highest number of sickle cell births by country. Available at: https://www.worldatlas.com/articles/countries-with-the-highest-number-of-sickle-cell-births-per-year.html. Accessed March 26, 2019.
18. López-Muñoz F, Shen WW, Pae CC, Moreno R, Rubio G, Molina JD, et al. Trends on literature on atypical antipsychotics in South Korea: A bibliometric study. Psychiatry Invest. 2013; 10: 8-16.
19. Okoroiwu HU, López-Muñoz F, Povedano-Montero FJ. Bibliometric analysis of global Lassa fever research (1970-2017): a 47-year study. BMC Infectious Disease. 2018; 18:639.
20. White HD, McCain KW. Blibliometric. Ann Rev Inf Sci Technol. 1989; 24: 119-186.
21. López-Muñoz F, Tracy DK, Povedano-Montero FJ, Breedvelt J, Garcia-Pacios J, Fernandez-Martin MP, et al. Trends in the scientific literature on atypical antipsychotic drugs in the United Kingdom: a bibliometric study. Therapeutic Advances in Psychology. 2019; 9: 1-2.
22. Falagas ME, Pitsouni EI, Melietsis GA, Pappas G. Comparison of Pubmed, Scopus, Web of Science and Google scholar: strength and weakness. FASEB J. 2008: 22(2): 338-342.
23. Kulkarni AV, Aziz B, Shams I, Busse JW. Comparisons of citations in Web of Science Scopus and Google Scholar for articles published in General Medical Journals. JAMA. 2009; 30(10): 1092-1096.
24. Povedano-Montero FJ, López-Muñoz F, Cruz FHS. Bibliometric analysis of the scientific production in the area of optometry. Arch Soc Esp Oftamol. 2016; 91(4): 160-169.
25. Price DJS. Little science, big science. New York. Columbia University Press. 1963.
26. Bradford SC. Documentation. London: Crosby Lockwood; 1948.
27. The World Bank. Available at: https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.25. Accessed March 26, 2019.
28. The World Health Organization. Available at: https://apps.who.int/nha/database/select/indicators/en. Accessed March 26, 2019.
29. Kapoor S, Little JA, Pecker LH. Advances in the treatment of sickle cell disease. Mayo Clin Proc. 2018; 93(12): 1810-1824.
30. López-Muñoz F, Vieta E, Rubio G, Garcia-Garcia P, Alamo C. Bipolar disorder as an emerging pathology in the scientific literature: a bibliometric approach. J Affect Disosrd. 2006; 92: 161-170.
31. Garcia-Garcia P, López-Muñoz F,, Callejo J, Martin-Agueda B, Alamo C. Evolution of Spanish scientific production in international obstetrics and gynecology journals during the period 1986-2002. Eur J Obstetr Gynecol. 2005; 123: 150-156.
32. López-Muñoz F, Alamo C, Quintero-Gutierrez FJ, et al. A bibliometric study of international scientific productivity in attention-deficit hyperactivity disorder covering the period 1980-2005. Eur Child

Adolesc Psychiatry. 2008; 17: 381-391.

33. National Institute of Health. Who we are. Available at: https://www.nih.gov/about_nih/who-we-are. Accessed March 29, 2019.
34. Institute national de la santé et la recherché medicale (INSERM). About Inserm. Available at: https://www.inserm.fr/en/about-inserm. Accessed March 29, 2019.
35. Small H. Co-citation in the scientific literature a new measure of the relationship between two documents. Journal of the American Society for Information Science. 1983; 24: 265-269.
36. Martyn J. Bibliographic coupling. Journal of Documentation. 1964; 20: 236.
37. Cancino CA, Merigo JM, Caronado FC. A bibliometric analysis of leading universities in innovation research. Journal of Innovation and Knowledge. 2017; 2: 106-124.
38. Vichinsky E. Chronic organ failure in adult sickle cell disease. Hematology Am Soc Hematol Educ Program/ 2017; 2017(1): 435-439.
39. Vichinsky E, Tones M, Minniti CP, Barrette S, Habr D, Zhang Y, et al. Efficacy and safety of deferasinox compared with deferoxamine in sickle cell disease: two-year results including pharmacokinetics and concomitant hydroxyurea. Am J Hematol. 2013; 88(12): 1068-1073.
40. Kassim A, DeBaun MR. The case for and against initiating either hydroxyurea therapy, blood transfusion therapy or hematopoietic stem cell transplant in asymptomatic children with sickle cell disease. Expert Opinion on Pharmacotherapy. 2014; 15(3): 325-336.
41. DeBaun MR, Kirkham FJ. Central nervous system complications in sickle cell disease. Blood. 2016; 127(7): 829-838.
42. Hanna J, Wernig M, Markoulaki S, Sun CW, Meissner A, Cassady JP, et al. Treatment of sickle cell anemia mouse model with ips cells generated from autologous skin. Science. 2007; 18(5858): 1920-1923.
43. Gao Y, Wang Y, Zhai X, He Y, Chen RE, Zhou J, et al. Publication trends of research on diabetes mellitus and T cells (1997-2016): A 20-years bibliometric study. PLOS ONE. 2017; 12(9): e0184869.
44. Li T, Ho YS. Bibliometric analysis on global Parkinson's disease research trends during 1991-2006. Neurosci Lett. 2008; 441: 248-252.
45. Zhao J, Yu G, Cai M, Lei X, Yang Y, Wang Q, et al. Bibliometric analysis of global scientific activity on umbilical cord mesenchymal stem cells: a swiftly expanding shifting focus. Stem Cell Research & Therapy. 2018; 9: 32.
46. Gómez I, Bordons M. Limitaciones en el uso de los indicadores bibliométricos para la evaluación científica. Política Científica.1996; 46:21-26.

Tables

Table 1: Document type

Document Type	No of documents	%
Article	12693	63.72
Review	3457	17.35
Letter	1170	5.87
Note	639	3.21
Conference Paper	555	2.79
Editorial	480	2.41
Book Chapter	429	2.15
Short Survey	302	1.52
Erratum	94	0.47
Article in Press	72	0.36
Book	25	0.13
Conference Review	5	0.03

Table 2: Bradford distribution of journals

	No of journals	% of journals	No of articles	% of articles	Bradford multiplier
Core	69	1.68	6672	33.49	
Zone 1	492	11.99	6548	32.87	7.13
Zone 2	3542	86.33	6701	33.64	7.19
Total	4103	100.00	19921	100.00	7.16

Table 3: Analysis of sources with the highest number of publications

	No of documents	Productivity Index	Impact Factor	Country of origin	Abbreviated Journal Title
Blood	642	3.22	15.132	USA	Blood
American Journal of Hematology	525	2.64	5.303	USA	Am. J. Hematol.
British Journal of Haematology	458	2.30	5.128	England	Br. J. Haematol.
Pediatric Blood and Cancer	360	1.81	2.642	USA	Pediatr. Blood Cancer
*Journal of Pediatric Hematology/Oncology**[1]**	243	1.22	1.060	Switzerland	Int. J. Pediatr. Hematol-Oncol.
Hemoglobin	239	1.20	0.462	USA	Hemoglobin
*Revista Brasileira de Hematologia e Hemoterapia**[2]**	193	0.97	0.62	Brasil	Rev. Bras. Hematol. Hemoter.
Transfusion	183	0.92	3.432	USA	Transfusion
Haematologica	182	0.91	9.090	Italy	Haematologica
PLoS ONE	168	0.84	2.766	USA	PLoS One

[1] Data from the *Journal Citation Report* (2017)

2 Data from 1999

3 Data from 2015

Table 4: Top 20 most productive countries.

Country	No of documents	%
United States	9167	46.02
United Kingdom	1869	9.38
France	1379	6.92
Brazil	933	4.68
India	660	3.31
Italy	577	2.90
Nigeria	574	2.88
Canada	522	2.62
Germany	429	2.15
Saudi Arabia	418	2.10
Netherlands	375	1.88
Turkey	307	1.54
Jamaica	232	1.16
Greece	216	1.08
Spain	215	1.08
Belgium	205	1.03
Australia	203	1.02
Switzerland	175	0.88
Egypt	150	0.75
Israel	140	0.70

Table 5: Comparative analysis of the productivity of the top 10 most productive countries in SCD research with their overall production in the area of medicine and other areas of biomedical research.

Country	IP Medicine	IP Health Professions	IP Pharmacology, Toxicology and Pharmaceutics	IP Sickle Cells	IP Tuberculosis	IP AIDS
United States	25.51	49.80	15.01	46.02	18.38	33.76
United Kingdom	8.07	14.43	4.14	9.38	8.04	7.43
France	3.84	0.66	1.71	6.92	4.50	3.72
Brazil	1.76	1.22	5.09	4.68	2.00	1.84
India	2.32	0.35	18.08	3.31	5.54	2.51
Italy	3.08	0.74	1.95	2.90	2.84	2.93
Nigeria	0.24	0.09	0.75	2.88	0.43	0.54
Canada	3.64	6.27	1.44	2.62	2.22	3.32
Germany	6.10	1.36	3.17	2.15	3.78	3.81
Saudi Arabia	0.41	0.06	0.62	2.10	0.50	0.15

Table 6: Top 20 most productive institutions

Institution	No of documents	%
National Institutes of Health, Bethesda	447	2.24
INSERM	396	1.99
The Children's Hospital of Philadelphia	350	1.76
Harvard Medical School	310	1.56
St. Jude Children Research Hospital	306	1.54
University College London (UCL)	294	1.48
UCSF Benioff Children's Hospital Oakland	290	1.46
Duke University School of Medicine	273	1.37
National Heart, Lung, and Blood Institute	268	1.35
The University of North Carolina at Chapel Hill	267	1.34
Albert Einstein College of Medicine of Yeshiva University	259	1.30
AP-HP Assistance Publique - Hopitaux de Paris	257	1.29
Children's Hospital Boston	243	1.22
King's College London	242	1.21
Hopital Henri Mondor	237	1.19
Cincinnati Children's Hospital Medical Center	235	1.18
The Johns Hopkins School of Medicine	220	1.10
Medical College of Wisconsin	219	1.10
University of Illinois at Chicago	217	1.09
University of Oxford	216	1.08

Table 7: Citation analysis of the 10 top most productive institutions.

Insttitution	No of documents	%	Citation	Citation/documents
National Institutes of Health, Bethesda	447	2.24	20253	45.31
Inserm	396	1.99	7306	18.45
The Children's Hospital of Philadelphia	350	1.76	11584	33.10
Harvard Medical School	310	1.56	14105	45.50
St. Jude Children Research Hospital	306	1.54	10181	33.27
UCL	294	1.48	13061	44.43
UCSF Benioff Children's Hospital Oakland	290	1.46	12856	44.33
Duke University School of Medicine	273	1.37	12499	45.78
National Heart, Lung, and Blood Institute	268	1.35	14187	52.94
The University of North Carolina at Chapel Hill	267	1.34	10983	41.13

Table 8: Ten most productive authors in SCD research

Author	Country	Affiliation	No of documents	% Documents	Citation	% Citacion
Vichinsky, E.	USA	UCSF Benioff Children Hospital	180	0.90	6713	1.91
DeBaun, M.R.	USA	Washington University School of Medicine	164	0.82	3463	0.99
Gladwin, M.T.	USA	Pittsburg Heart, Lung, Blood and Vascular Medical Institute	163	0.82	10793	3.08
Steinberg, M.H.	USA	Pittsburg Heart, Lung, Blood and Vascular Medical Institute	157	0.79	6062	1.73
Ballas, S.K.	USA	Cardeza Foundation for Hematologic Research	155	0.78	4039	1.15
Galactéros, F.	France	Henri-Mondor Hospital	147	0.74	3106	0.89
Ware, R.E.	USA	Cincinnati Children’s Hospital Medical Centre	146	0.73	5556	1.58
Kato, G.J.	USA	Pittsburg Heart, Lung, Blood and Vascular Medical Institute	145	0.73	5239	1.49
Wang, W.C.	USA	St. Jude Children Research Hospital	138	0.69	6692	1.91
Connes, P.	France	Universite Claude Bernard Lyon 1 Villeubanne; LABEX GR-Ex; Institut universitaire de France	114	0.57	1341	0.38

Table 9: Top 10 most cited articles in SCD research

Article	Authors	Source	Cited	% Citacion
Heart disease and stroke statistics-2012 update: A report from the American heart association	Roger, V.L., Go, A.S., Lloyd-Jones, D.M., (...), Woo, D., Turner, M.B.	*Circulation*	3511	1.00
Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013	Naghavi, M., Wang, H., Lozano, R., (...), Sabin, N., Temesgen, A.M.	*The Lancet*	2055	0.59
Glutathione Metabolism and Its Implications for Health	Wu, G., Fang, Y.-Z., Yang, S., Lupton, J.R., Turner, N.D.	*Journal of Nutrition*	1969	0.56
Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013	Vos, T., Barber, R.M., Bell, B., (...), Salomon, J.A., Murray, C.J.L.	*The Lancet*	1523	0.43
Guidelines for the prevention of stroke in patients with stroke or transient ischemic attack: A guideline for healthcare professionals from the American Heart Association/American stroke association	Furie, K.L., Kasner, S.E., Adams, R.J., (...), Turan, T.N., Wentworth, D.	*Stroke*	1160	0.33
Updated clinical classification of pulmonary hypertension	Simonneau, G., Gatzoulis, M.A., Adatia, I., (...), Robbins, I.M., Souza, R.	*Journal of the American College of Cardiology*	1143	0.33
Guidelines for the prevention of stroke in patients with stroke and transient ischemic attack: A guideline for healthcare professionals from the American Heart Association/American Stroke Association	Kernan, W.N., Ovbiagele, B., Black, H.R., (...), Schwamm, L.H., Wilson, J.A.	*Stroke*	1140	0.32
Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin	Hanna, J., Wernig, M., Markoulaki, S., (...), Townes, T.M., Jaenisch, R.	*Science*	1097	0.31
Guidelines for prevention of stroke in patients with ischemic stroke or transient ischemic attack: A statement for healthcare professionals from the American Heart Association/American Stroke Association council on stroke - Co-sponsored by the council on	Sacco, R.L., Adams, R., Albers, G., (...), Schwamm, L.H., Tomsick, T.	*Stroke*	1086	0.31

cardiovascular radiology and intervention. The American Academy of Neurology affirms the value of this guideline					
Hematopoietic stem-cell transplantation	Copelan, E.A.	*New England Journal of Medicine*	1058	0.30	

Figures

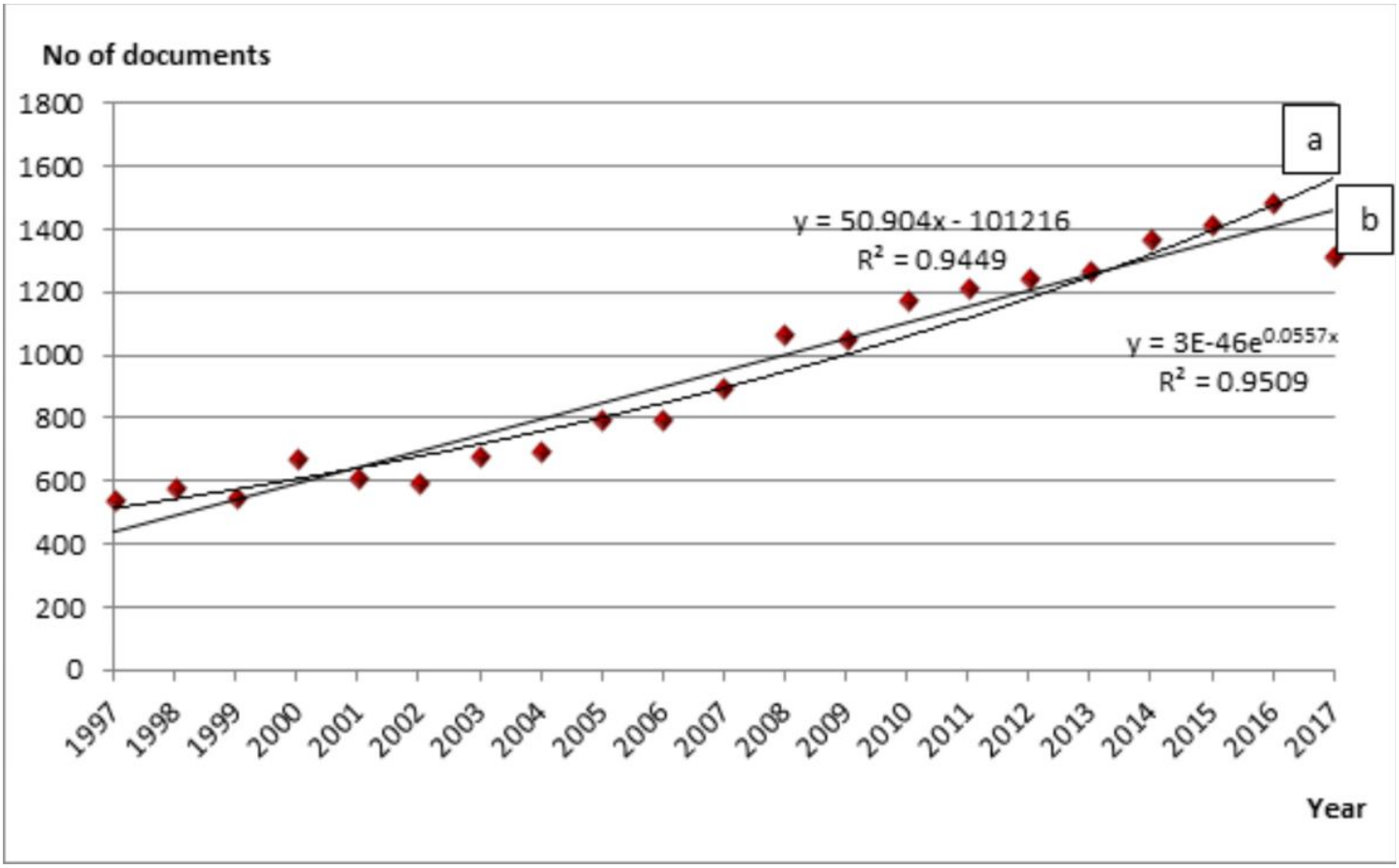

Figure 1

Chronological distribution of scientific literature on sickle cell disease research within the study period. (a) Linear trendline. (b) Exponential trendline.

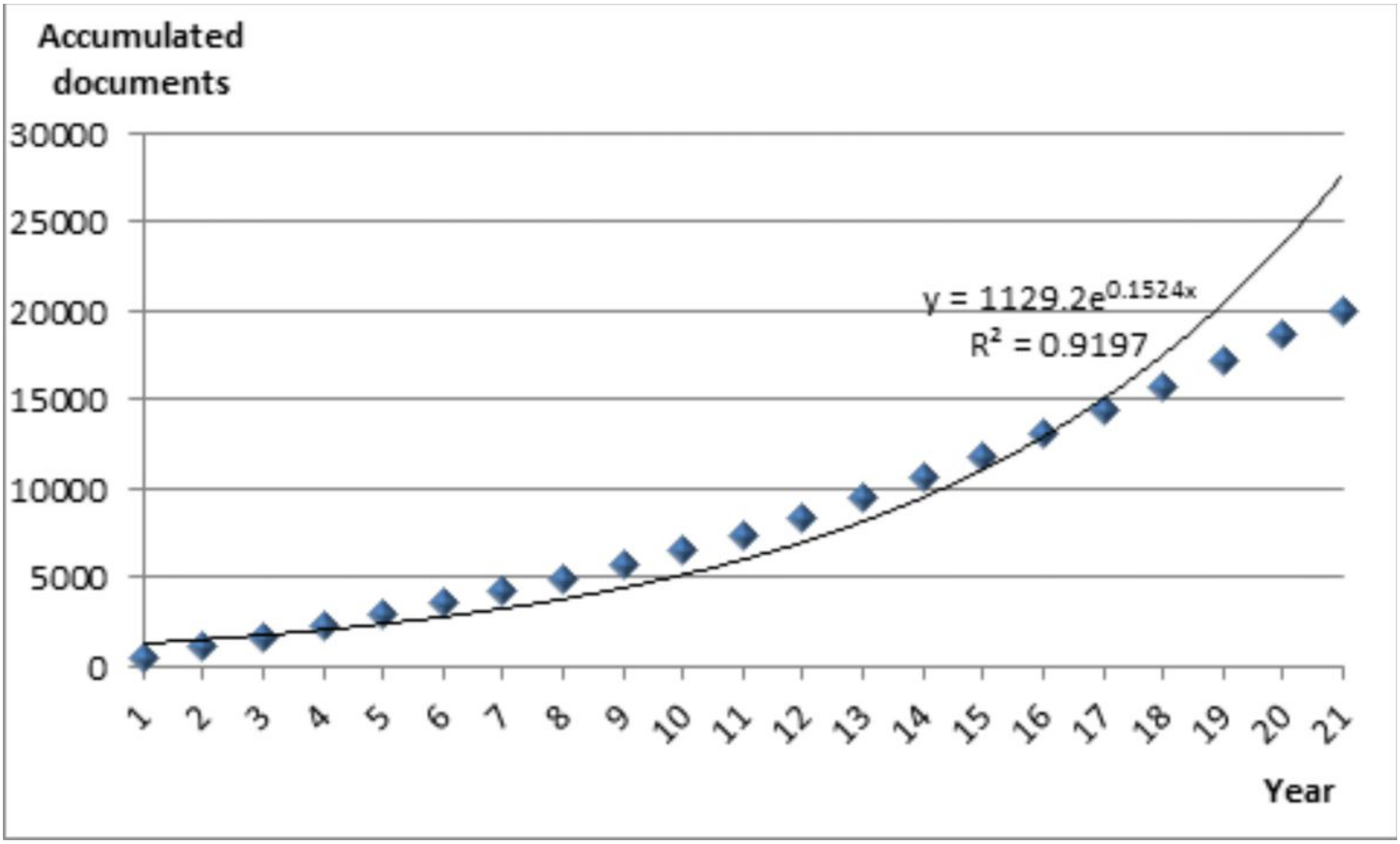

Figure 2

Temporary evolution of publications on sickle cell disease. Production doubles every 4.52 years. (formula can be found in the supplemental files

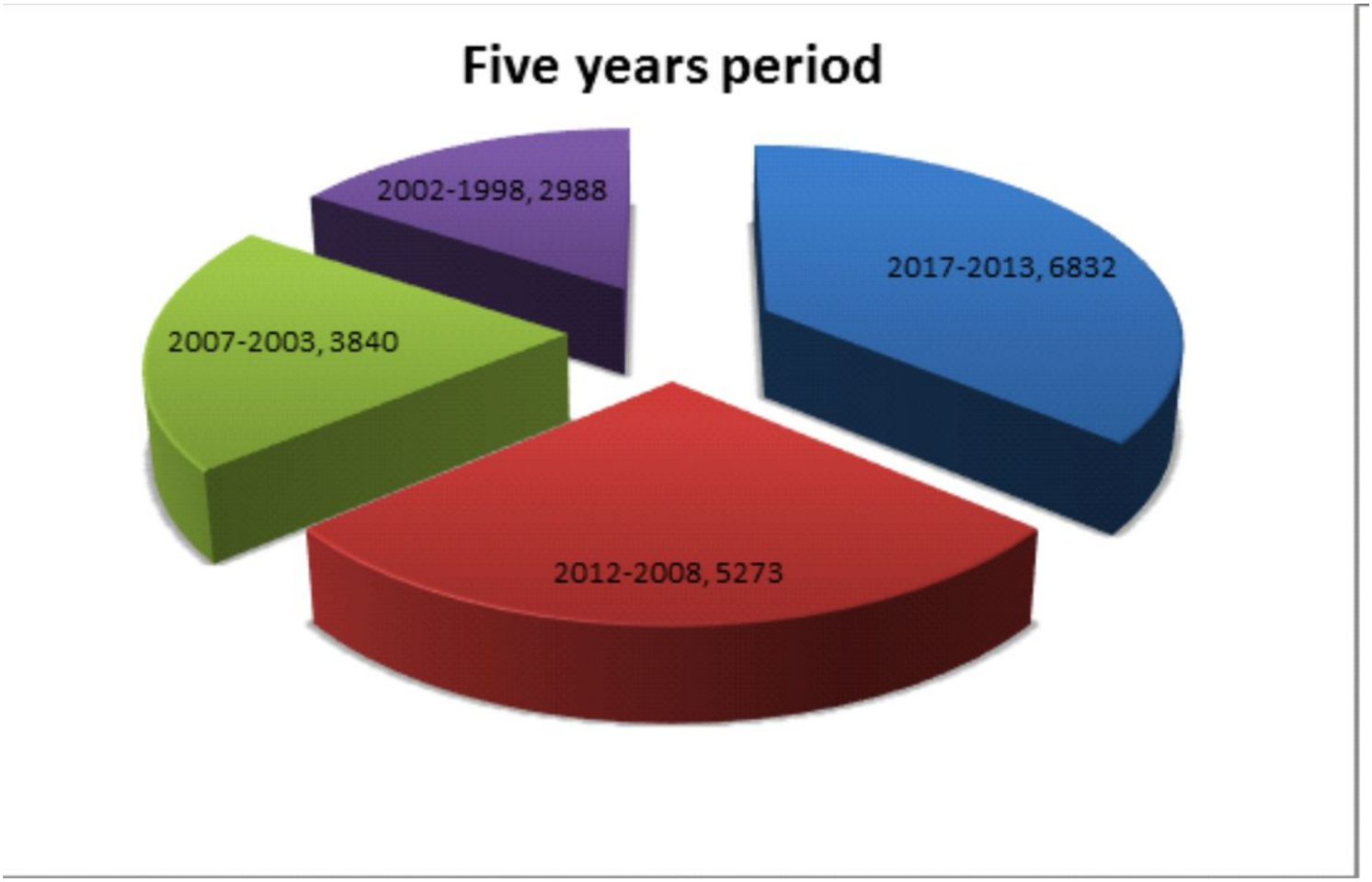

Figure 3

Cumulative growth in total global sickle cell disease research in each 5 – year period.

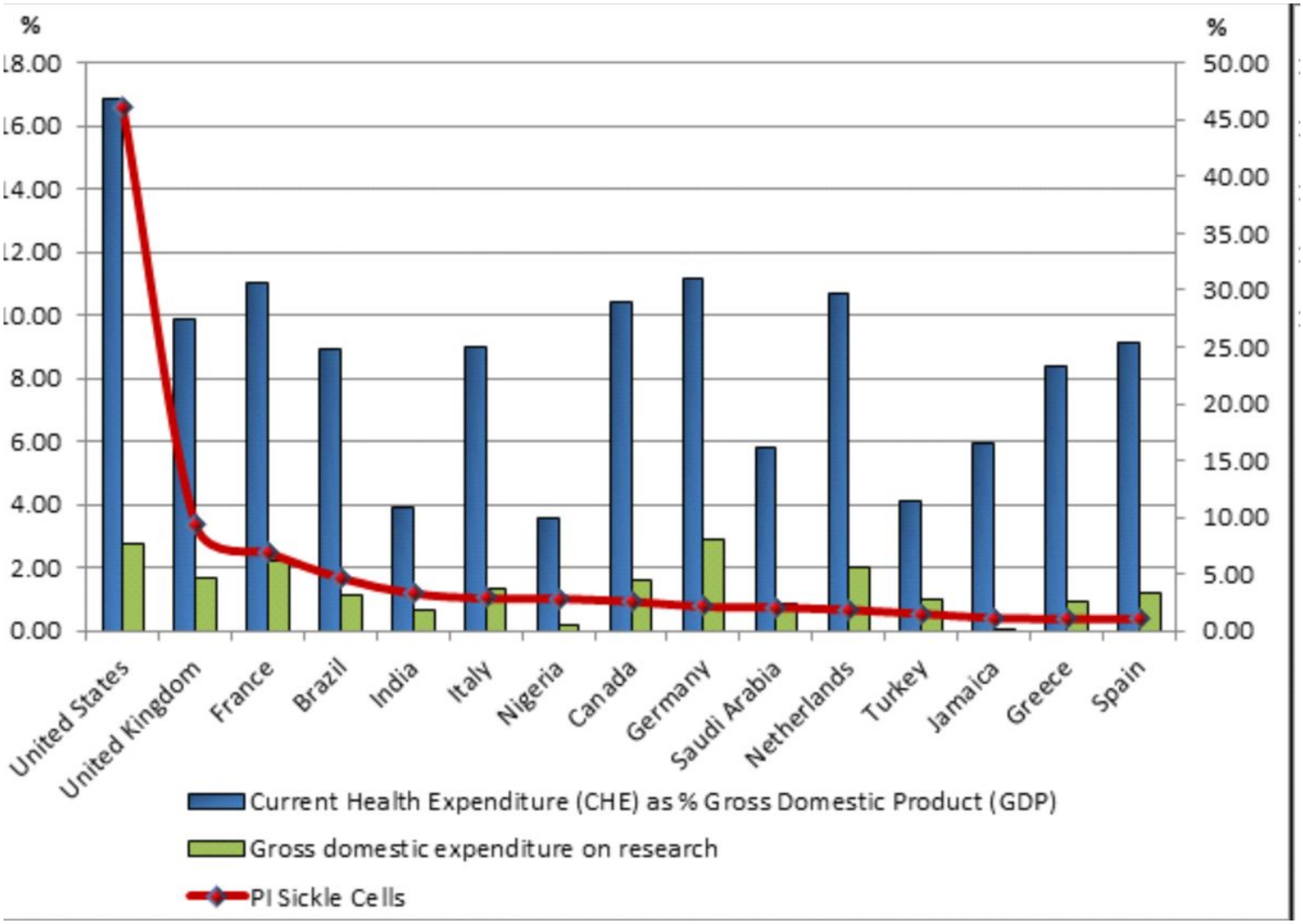

Figure 4

Relationship between production of scientific literature on SCD and current health expenditure (CHE) as % GDP and Gross Domestic Expenditure on Research and Development (R&D) in the top 15 most productive countries on SCD research.

Figure 5

Bibliometric coupling among most productive institutions in SCD research.

Figure 6

Keyword analysis among the articles in the studied repertoire.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Figure2formula.PNG
- Formula1.PNG