Machine learning in healthcare strategic management: a systematic literature review

Sand Mohammad Salhout
Graduate Studies, Arab American University, Ramallah, Palestinian Authority

Received 25 June 2023
Revised 1 November 2023
Accepted 14 November 2023

Abstract

Purpose – This study specifically seeks to investigate the strategic implementation of machine learning (ML) algorithms and techniques in healthcare institutions to enhance innovation management in healthcare settings.

Design/methodology/approach – The papers from 2011 to 2021 were considered following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. First, relevant keywords were identified, and screening was performed. Bibliometric analysis was performed. One hundred twenty-three relevant documents that passed the eligibility criteria were finalized.

Findings – Overall, the annual scientific production section results reveal that ML in the healthcare sector is growing significantly. Performing bibliometric analysis has helped find unexplored areas; understand the trend of scientific publication; and categorize topics based on emerging, trending and essential. The paper discovers the influential authors, sources, countries and ML and healthcare management keywords.

Research limitations/implications – The study helps understand various applications of ML in healthcare institutions, such as the use of Internet of Things in healthcare, the prediction of disease, finding the seriousness of a case, natural language processing, speech and language-based classification, etc. This analysis would help future researchers and developers target the healthcare sector areas that are likely to grow in the coming future.

Practical implications – The study highlights the potential for ML to enhance medical support within healthcare institutions. It suggests that regression algorithms are particularly promising for this purpose. Hospital management can leverage time series ML algorithms to estimate the number of incoming patients, thus increasing hospital availability and optimizing resource allocation. ML has been instrumental in the development of these systems. By embracing telemedicine and remote monitoring, healthcare management can facilitate the creation of online patient surveillance and monitoring systems, allowing for early medical intervention and ultimately improving the efficiency and effectiveness of medical services.

Originality/value – By offering a comprehensive panorama of ML's integration within healthcare institutions, this study underscores the pivotal role of innovation management in healthcare. The findings contribute to a holistic understanding of ML's applications in healthcare and emphasize their potential to transform and optimize healthcare delivery.

Keywords Systematic literature review, Machine learning, Healthcare institution

Paper type Literature review

1. Introduction

Machine learning (ML), a part of artificial intelligence (AI), studies computer algorithms that improve over time and support an event's prediction. It mainly comprises four types: supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. ML algorithms have been widely used in various applications such as social networking, virtual personal assistance, the banking sector, industry automation, self-driving cars, the healthcare industry, etc. ML algorithms are gaining immense popularity in

Arab Gulf Journal of Scientific Research
Emerald Publishing Limited
e-ISSN: 2536-0051
p-ISSN: 1985-9899
DOI 10.1108/AGJSR-06-2023-0252

medical diagnosis, hospital management and the healthcare sector. Many countries have allocated huge investments toward the healthcare sector.

Indian Brand Equity Foundation (Pervez, Dutta, & Parihar, 2022), the government of India has decided to increase the healthcare market threefold trillion (US$ 133.44 billion). According to a report by the Indian Brand Equity Foundation (IBEF), healthcare is one of the largest sectors in the country. According to the Centers for Medicare and Medicaid Services (Dugan & Spetz, 2022) (CMS), National health spending is projected to grow at an average annual rate of 5.4% from 2019 to 2028 and to reach $6.2 trillion by 2028. The health share of the economy is projected to rise from 17.7% in 2018 to 19.7 percent in 2028.

Health care is a broad umbrella that includes hospitals, medical devices, clinical trials, outsourcing, telemedicine, medical tourism, health insurance and medical equipment. Managing these is challenging, and two of the most difficult jobs are hospital bed management and patient footfall management. Managing them can reduce the waiting time and increase the availability of hospitals. The primary reason is that ML and AI algorithms can play an essential role in advancing the healthcare sector by analyzing massive datasets and extracting hidden information and patterns from the datasets to predict the outcomes eventually.

To address these problems, ML and AI come into the picture. Supervised or unsupervised, deep learning (DL) models can be applied for disease prediction and diagnosis. Managing hospital resources is one of the significant tasks in the healthcare industry; time series analysis can be used for estimating future outcomes where the previous history can be analyzed, and actions related to the future can be taken. Much work has been done in disease prediction and developing monitoring applications for various diseases.

Coccia (2018) looked into the emergence of new research fields; ML would qualify as such. The author identified four properties; for example, a new discipline can originate from a process of specialization within applied sciences, as the case with ML.

In the expansion of the properties mentioned above, Coccia (2018, 2020a) investigated the evolution of experimental physics. The article analyzed the structure and processes of the properties of the evolution of science (with a focus on experimental physics).

Coccia (2020c) analyzed disruptive technologies such as ML, their origin and their impact on the public and private sectors.

Mueller *et al.* (2023) looked at health impact assessment and identified seven relevant models, among them ML.

Gad *et al.* (2023) demonstrated that breast cancer diagnoses can be improved by applying ML and swarm intelligence. The achieved accuracy was above 97%. Zhang *et al.* (2023) also looked into medical imaging techniques to diagnose breast cancer, especially the predictive and prognostic features of breast cancer by radiomics.

Coccia (2020b) stressed the socioeconomic benefits of ML for poorer regions because these can send digital images for diagnoses to labs in developed countries, thereby reducing the current gap in healthcare services.

Mohammadpour *et al.* (2022) analyzed Clinical Decision Support Systems, which were divided into two groups: rule-based (clinical guidelines) and non-rule-based (ML). Their assessment showed that Rule-based and non-rule-based-decision support systems have different benefits and should be combined, especially in cancer diagnosis.

Chhetri, Goyal, and Mittal (2023) looked at ML in the context of psychological disorders. Although only based on a review of 26 articles, the results show that ML can help in brain imaging, behavioral kinematics and memory analysis.

Oladimeji (2023) conducted a bibliometric analysis on intelligent health research and found that the journal IEEE Access has the highest number of papers in this discipline, and most publications originated from India.

This paper conducts a systematic literature review to understand the trends and applications of ML in hospital management. The author considers various hospital management fields such as clinical study, medical history, hospital readmission, length of stay, etc. "Scopus," the most extensive database, is used to find relevant documents in this area. A bibliometric literature review is used to lighten the emerging areas in this area. Large volumes of scientific data are analyzed and explored to do so. English language articles with open access, which are at the final stage, have been filtered out for analysis. The objective of this paper is based on the following research questions:

(1) What are the prevalent ML techniques and approaches used in healthcare research, particularly in the context of diseases like hypertension, diabetes mellitus, cerebrovascular disease, cardiovascular disease, acute kidney failure, heart failure, cancer diagnosis and chronic obstructive lung disease?

(2) How does the application of ML in healthcare management intersect with various medical disciplines, and what insights can be gained from this interdisciplinary perspective?

(3) What are the potential benefits and impacts of a holistic approach to ML in healthcare management, and how does it improve overall healthcare outcomes?

(4) Which countries, publications and authors have been at the forefront of research in ML applied to healthcare?

The paper follows the following sequence: Section 1 explains the critical peer literature review, followed by Section 2, conducting a bibliometric analysis of the articles. Results and findings from the previous section are presented in Section 3 of the study. Section 4 presents discusses the study and provides a comparative analysis of our research about existing works; finally, Section 5 deals with the paper's conclusion with limitations and future suggestions.

2. Theoretical reference framework

To achieve the objective of performing a systematic literature review on ML in healthcare sector management, various related works were studied. The study of multiple methodologies applied by the existing papers helped formulate the research methodology. The study on related jobs also helped gain which analyses should be performed.

AI is the broadest field focused on creating intelligent machines. ML is a subset of AI that emphasizes learning from data, and DL is a subset of ML that uses deep neural networks (Shams & Kadow, 2023). In healthcare, these technologies work together to enhance diagnostics, treatment and healthcare system management. ML is a subset of AI that focuses on training algorithms to learn from data and make predictions or decisions based on that learning. It's a data-driven approach to AI, where algorithms improve their performance over time as they process more data. DL, on the other hand, is a subset of ML that employs explicitly artificial neural networks inspired by the structure and function of the human brain. DL has gained attention due to its success in complex tasks, such as image and speech recognition.

Bolhasani, Mohseni, and Rahmani (2021), in their work of conducting a systematic literature review on DL in the healthcare sector, collected documents from various electronic databases: IEEE, Springer, Google Scholar, ACM and Elsevier. They included the only English papers with more than 20 citations in 2010–2020 and obtained 44 documents. From the study on the records obtained from the papers, the researchers found various frameworks, platforms, Internet of Things (IoT) technologies, and DL algorithms used in medical diagnosis, disease prediction and home and personal healthcare-based applications.

A structured literature review (SLR) was performed by Born *et al.* (2021) on the role of AI in medical imaging for COVID-19 disease. They adopted the standard Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) method for finding relevant papers. To perform a meta-analysis, they assigned each manuscript a score from 0 to 1 to determine the maturity level. They concluded in their study that for a robust AI solution to be developed, it is essential to improve interdisciplinary collaborations.

Natural language processing (NLP) is an essential field of ML. Its application in healthcare is significantly growing, according to Casey *et al.* (2021). In radiology, i.e., NLP can extract meaningful information from a medical report. Although DL models are still used, the authors found that ML algorithms provide better results. One of the significant drawbacks of DL models is that they mostly use a black-box approach, making them difficult to understand and apply in the medical field.

Secinaro, Calandra, Secinaro, Muthurangu, and Biancone (2021) aimed to determine the overall role of AI in the healthcare sector. They followed the SLR process and obtained the most relevant document from the Scopus database, and a bibliometric analysis was performed. The results revealed that AI in healthcare is an emerging field, and most people are concentrating on health services management, predictive medicine, patient data and diagnostics and clinical decision-making. The United States, United Kingdom and China researchers have highly contributed to this field.

Speech and language classification is one of the essential parts of ML. It has various applications in the field of management and customer segmentation. Vigo, Coelho, and Reis (2022) aimed to discover the practical algorithms and best practices for Alzheimer's disease. The authors could find the essential resources and data models required for a decision support system for Alzheimer's disease and a list of speech features correlated with the disease's linguistic and acoustic footprint.

Emergency department (ED) management is required in the healthcare sector domain. Due to the advent of the COVID-19 disease, its importance has grown exponentially. In their work, Ortíz-Barrios, Coba-Blanco, Alfaro-Saíz, and Stand-González (2021) performed a systematic review of how the process can be improved to increase the effective response of the ED. They revealed that ML and data analytics effectively predict COVID-19. They recommended integrating AI in operations and quality management sections to improve the response of ED. Raj and Manivannan (2022) proposed ML models for predicting COVID-19 vaccination intention. They considered an Indian market for analysis purposes.

AI-based technologies can help aid and improve care (Von Gerich *et al.*, 2022). The keywords used were related to AI, nursing and ML algorithms. Their study concluded that most documents published were in operations, testing, implementation and descriptive design. Saleem, Al-ghamdi, Alassafi, and Alghamdi (2022) worked to determine how effective ML, DL and mathematical models are in forecasting and classifying COVID-19 disease. An SLR was performed following PRISMA guidelines. They found that the convolutional neural network (CNN), a DL model and the support vector machine (SVM) ML model outperformed others in predicting and classifying COVID-19 disease.

The use of current clinical data can improve prognosis. Xu, Wang, and Sansgiry (2020) conducted a systematic literature review on predicting diabetic retinopathy, nephropathy and neuropathy in patients with type 1 diabetes using ML. They also considered a cross-reference for finding relevant documents for the study. They found that age, duration of diabetes, BMI, BP and lipids of current HbA1C value are the prominent factors to be used for the prediction purpose. Classification and Regression Trees (CART), Random Forest (RF), Support Vector Machine (SVM), Logistic Regression (LR) and Convolutional Neural Network (NN) were used for prediction.

Giannakopoulou and Roussaki (2022) aimed to discover the prominent IoT and ML methods for monitoring and managing Parkinson's disease. SLR was conducted, and they found that inertial, pressure and force, image and depth, and voice and audio sensors were

most commonly used. CNN was found to be the best model as per the ranking results. AI techniques are widely used to predict and diagnose cancer (Kumar, Gupta, Singla, & Hu, 2021). The authors performed SLR and concluded that tobacco, excess weight and obesity account for 53% of the total causes of cancer. They also found that various techniques are to be employed for multiple types of cancer.

Saxena, Deogaonkar, Pais, and Pais (2023) the authors examined the connection between feelings and job efficiency following COVID-19. Seventy-two survey participants from a mid-sized consulting firm that participated in the study provided text inputs, and these sentiments were collected and correlated with the productivity scores. The sentiment score was calculated using a machine-learning model created in Python.

AI techniques, ranging from ML to DL, can provide remote assistance to patients (Kumar, Koul, Singla, & Ijaz, 2022). The researchers conducted an overall review of how AI is used for treating various diseases. The results section mentioned how many papers are being published for each condition. Which algorithm is widely being used for a particular disease is also mentioned.

Coccia (2020b) examined DL applications in cancer imaging, illuminating key critical developments in the industry. It demonstrated how DL may change cancer diagnoses and even eliminate regional disparities in healthcare. Case studies in the research illustrated how well DL detected lung cancer and evaluated breast cancer lymph node metastases. The study demonstrated the potential for more accurate cancer diagnosis through the combination of DL and modern computing technologies, providing a way to lessen invasiveness and shorten diagnostic wait times. In addition to improving patient outcomes, DL reduces the financial burden on healthcare systems. The study examined barriers to incorporating DL into clinical practice while acknowledging its promise and offering suggestions for overcoming them.

Findings from EEG, brain imaging, behavioral and kinematic investigations on ML applications in addiction research were compiled by Chhetri *et al.* (2023). The key conclusions indicated that diagnosing drug-related diseases requires a disease-based strategy. Psychoactive chemicals have an impact on brain function, and the behavioral alterations linked to substance use are the most startling discoveries. To categorize, detect and forecast addictions, ML algorithms have recently acquired popularity in the field of addiction. Planning, monitoring and treatment management can lessen the burden of disease caused by addiction in addition to technology-based therapies.

The development, use and challenges of radionics and AI in detecting breast cancer were covered by Zhang *et al.* (2023). The study's findings indicate a significant discrepancy between research findings and their clinical use. Modern DL approaches have not been widely used in the sector, partly because most research only used small datasets. Future radionics technologies powered by AI may be incorporated into clinical practice to improve the accuracy of breast cancer diagnosis. To choose the best course of clinical practice, this integration would combine imaging data with patient information.

From the study of related works above, our job is different as it concentrates on understanding an overall implementation of ML in healthcare institutions' management rather than just focusing on a particular disease or area of healthcare management. Compared with other related works, the paper considers various study regions, namely medicine, biochemistry, genetics, nursing, etc. The paper uses the R software bibliometrics package to perform bibliometric analysis.

This approach helps get an overall picture of how effective ML algorithms work in healthcare management. In addition, unlike some papers that considered only a short study timeline for performing SLR, this paper considers the records of ten years from 2011 to 2021. This would allow the future researcher to target the most flourishing areas of the healthcare sector using ML techniques with great scope in the future. Table 1 presents the comparison of field of study of recent systematic literature reviews on the healthcare sector, pertaining to the field explored in this study.

References	Field of study	Period of analysis
Bolhasani *et al.* (2021)	Deep learning applications for IoT in healthcare	2010–2020
Born *et al.* (2021)	Artificial intelligence in medical imaging of Covid-19	2017–2020
Casey *et al.* (2021)	Application of natural language processing to radiology reports	2015–2019
Castelyn, Laranjo, Schreier, and Gallego (2021)	Predictive performance and impact of algorithms in remote monitoring of chronic conditions	2013–2018
Di Tanna, Wirtz, Burrows, and Globe (2020)	Risk prediction models for adults with heart failure	2013–2018
Fernandes, Paredes, and Salinet (2020)	Cardiac Arrhythmias detection and classification by machine learning	2014–2019
Giannakopoulou and Roussaki (2022)	IoT and ML for monitoring, management and diagnosis of Parkinson's disease	2012–2021
Kumar *et al.* (2021)	Artificial intelligence techniques in cancer prediction and diagnosis	2009–2021
Kumar *et al.* (2022)	Artificial intelligence in disease diagnosis	2009–2020
Ortíz-Barrios *et al.* (2021)	Process improvement approaches for increasing the response of emergency departments against the COVID-19 pandemic	2019–2021
Pashazadeh and Navimipour (2018)	Extensive data handling mechanisms in healthcare applications	2006–2016
Saleem *et al.* (2022)	Machine learning, deep learning and mathematical models to analyze forecasting and epidemiology of COVID-19	2020–2021
Secinaro *et al.* (2021)	Artificial intelligence in healthcare	1992–2021
Vigo *et al.* (2022)	Speech- and language-based classification of Alzheimer's disease	2015–2020
Von Gerich *et al.* (2022)	Artificial intelligence-based technologies in nursing	2010–2021
Xu *et al.* (2020)	Predicting diabetic retinopathy, nephropathy and neuropathy in patients with type 1 diabetes using machine learning	2016–2019
This paper	Machine learning in the management of healthcare institutions	2011–2021

Source(s): Prepared by the Author (2023)

Table 1.
List of related work using systematic literature review on the healthcare sector

The investigation offers the following insights for future researchers and practitioners:

(1) In total, 123 peer-reviewed articles bibliometric analysis from Scopus Database.

(2) Information on co-occurrence analysis using All keywords using the R software bibliometrics package.

(3) Information on Co-occurrence analysis using Author keywords using the R software bibliometrics package.

(4) Performing citation analysis based on sources using R software bibliometrics package.

(5) Performing citation analysis based on countries using R software bibliometrics package.

3. Methodology

The paper aims to understand the applications and use of ML in managing healthcare institutions following an SLR methodology. The Systematic Literature Review deals with

collecting secondary data followed by performing analysis on them to answer a stated question. It helps gain insights on relevant authors, keywords, number of publications per year, sources, countries, coauthors, etc., related to a particular area of interest. SLR provides an intense basement for future researchers interested in a specific site.

The entire methodology is divided into three stages, namely.

(1) Sample and data

(2) Measures of variables

(3) Data analysis procedure

In summary, the researcher follows the following strategy, as shown in Table 2 to perform a Systematic Literature Review where keywords are identified first, then relevant documents based on the keywords. In the third step, screening of the process, obtained documents are filtered based on the requirements. Finally, the filtered documents are checked by eligibility to confirm their relevance to the raised question.

The proposed method will provide relevant records at the end, which will be further used in performing bibliometric analysis. Figure 1 depicts the proposed methodology for the study.

Sample and data: In the starting section, 'the required keywords are identified. The keywords are mainly related to various topics of ML, such as NLP, supervised and unsupervised education, computer vision, neural networks, reinforcement learning, predictive and time series analysis. At the same time, the other keywords were related to therapeutic areas, hospital management and healthcare institutions. Once the keywords were finalized, the query was formulated in the Scopus database. The study only concentrates on the past ten years (2011–2021) to see the trend of ML in the healthcare sector. The results obtained comprise 4680 documents that were relevant to the keywords and period mentioned. No databases other than the Scopus database were used to identify other records.

Measures of variables: In the second step of the proposed methodology, the identified records are checked for duplication. The screening process is performed to obtain a more relevant set of records. It is a data reduction step performed to reduce the number of forms.

Review protocol elements	Authors' consideration	Results in terms of sources
What is already known? And research topics	Machine learning (ML) has become a buzzword in the market from already known knowledge; it has many applications, including healthcare management. Due to the popularity and significant growth of ML, there is the potential for a structured literature review (SLR) investigating how ML can contribute to healthcare management	4,680
Journals' and thematic limitations	The journals related to various fields, namely medicine, nursing, biochemistry, health professions, immunology, pharmacology, toxicology, pharmaceutics, and many more were considered. Thus, the study provides a broader picture of ML applications in various healthcare sectors rather than confining them to a particular disease or field	2,133
Other restrictive elements	The authors selected only peer-reviewed articles, excluding conference proceedings, books, reviews, and book chapters. Finally, only sources written in English are considered	123
Period of analysis	2011–2021	

Source(s): Prepared by the Author (2023)

Table 2. SLR protocol and results obtained from Scopus

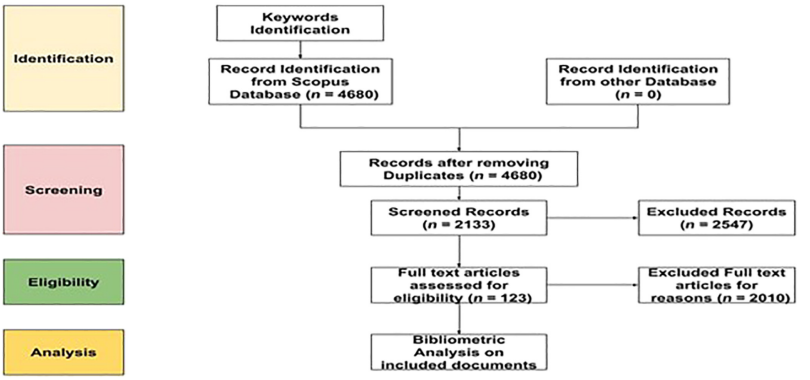


Figure 1.
Proposed methodology of the study

Source(s): Secinaro *et al.* (2021)

The fewer the records and documents, documents and more efficient for analysis. Two thousand one hundred thirty-three records were obtained from the screening step and, the obtained set of documents from the previous step is filtered. The following filters were applied to the documents obtained from the screening stage:

(1) A total of 75 keywords were selected.

(2) Only articles were selected.

(3) Open-access articles were chosen.

(4) Final stage.

(5) English language.

As a result of the above filter set, 123 documents were included.

Data analysis procedure: In the final step of the methodology, "Bibliometric analyses" were performed on the selected 123 documents obtained from the "measures of variables" stage. The study concentrates on gaining insights on:

(1) Co-occurrence analysis using All keywords.

(2) Co-occurrence analysis using Author keywords.

(3) Citation analysis based on sources.

(4) Citation analysis based on countries.

Performing the above analysis will help gain insights into the prominent keywords of this study, which country has done more work in this area, and which journal has more citations in this particular field of study.

4. Results

R studio's "Bibliometrics package" was used to conduct research. After obtaining relevant records from the Scopus Database, bibliometric analysis was performed to gain insights into them. First, a detailed table was created containing vital bibliometric statistics. Then, the analysis was performed to answer which country has the highest number of papers related to the study, which are the top sources with the highest number of citations, the essential keywords related to the survey, etc.

Bibliometrics is an R software package that provides all the instruments for bibliometric analysis. Bibliophagy 4.0 web application of the Bibliometrics package performs a comprehensive science mapping analysis on various databases such as Scopus, Web of Science, PubMed, etc. The application provided insights into different categories such as sources, authors, documents, conceptual structure and social structure. The bibliometric analysis can be divided into two major categories: level of analysis and K structures.

The level of analysis mainly focuses on the domain, which includes information on sources, authors and documents using different metrics. Other information, such as primary data and annual scientific production, is also provided. On the other hand, K structures focus on knowledge structure, including conceptual, intellectual and social information, using different bibliometric techniques. Table A1 in Appendix 1 shows the vital bibliometric statistics.

Document per author = Documents/authors.

Coauthors per documents = Authors appearances/documents.

Coauthors per documents = Authors appearances/documents.

Collaboration index = Authors of multiauthored articles/multiauthored articles.

4.1 Overview analysis

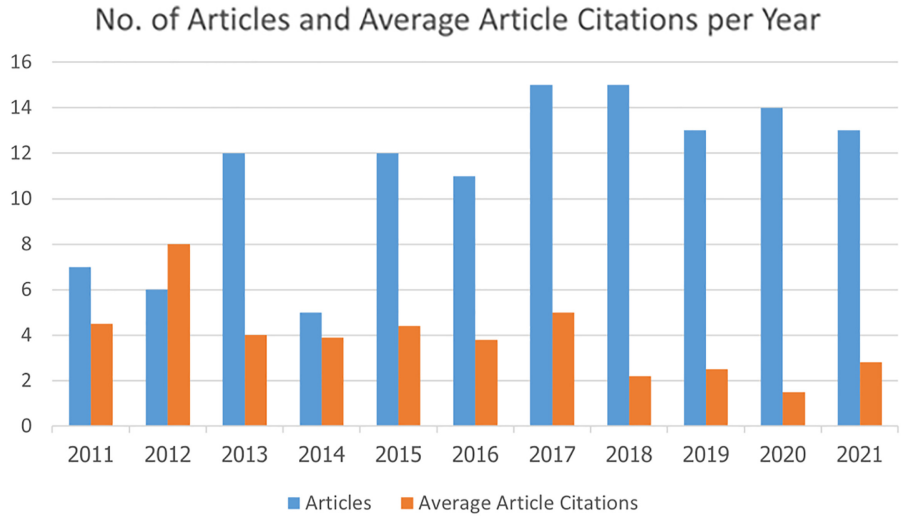
The graph (Figure 2) reveals that the annual scientific production steadily increased from 2011 to 2021. This also represents the average article citations per year. It was found that the highest number of sources was obtained for one or more articles published in 2012.

Figure 3 illustrates the technologies applied in the integration of DL in healthcare, with a focus on prevalent diseases.

4.2 Sources analysis

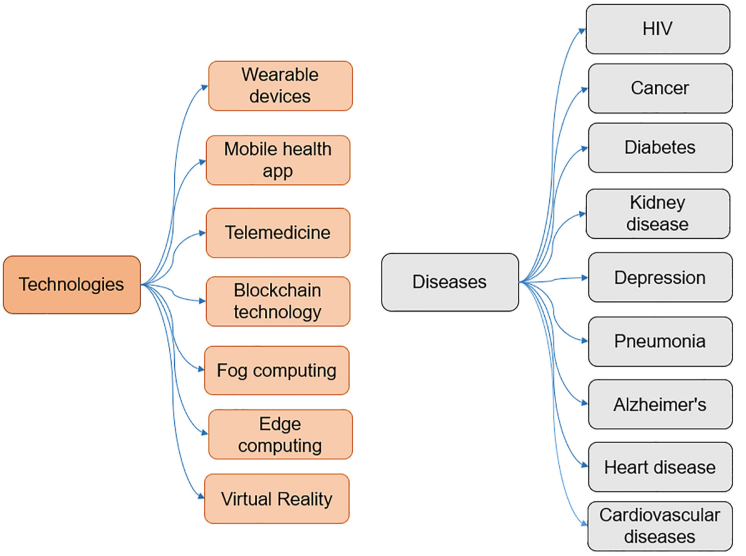
The source is a journal/book/conference proceeding series, etc., which publishes one or more documents in our bibliographic collection. From the analysis, it was found that there are 92 sources related to our collection. Figure 4 shows the most relevant sources.

For the publication of an article or to get good information about a subject, it is essential to know the top journals in that area. The above bar chart represents the top ten most



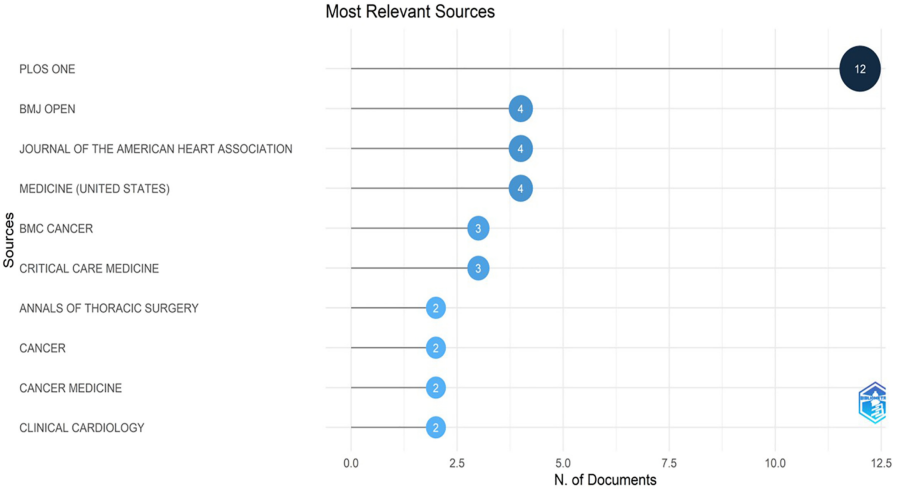
Source(s): Prepared by the author (2023)

Figure 2.
Number of articles and average article citations per year



Source(s): Prepared by the author (2023)

Figure 3.
Graph showcasing the most relevant technologies and diseases



Source(s): Prepared by the author (2023)

Figure 4.
Graph showcasing the most relevant sources

relevant sources of ML in the healthcare sector. It is depicted that *PLoS One* is the most pertinent journal in this area, with twelve published articles. *BMJ Open*, *Journal of the American Association* and *Medicine* (United States) journals also have relevant reports on the study.

The cited source is a source cited by one or more documents. In simpler terms, it is either a journal/book/conference proceeding, etc., which is added to the reference list by at least one of the filtered documents from the Scopus Database. It is found that 1667 sources are included in

123 documents' bibliographies. Figures A1 and A2 in Appendix 2 demonstrate mostly cited sources and source clustering.

H index, g index and m index values were calculated (Table 3) to gain insights into the most relevant sources. The values obtained after the analysis showed that *PLoS One* journal has the highest source impact for all three indexes.

H index: The Hirsch index is the number of articles published by an author or journal, each of which is cited in other papers at least once.

M index: It is simply the H index divided by the number of years since the first published paper of the journal.

G index: Improve the h index to measure the global citation performance of a set of articles. See Appendix 3 for more details.

Figure 5 depicts the source number of publications per year. It can be seen that there is an increase in the growth of almost all the top sources.

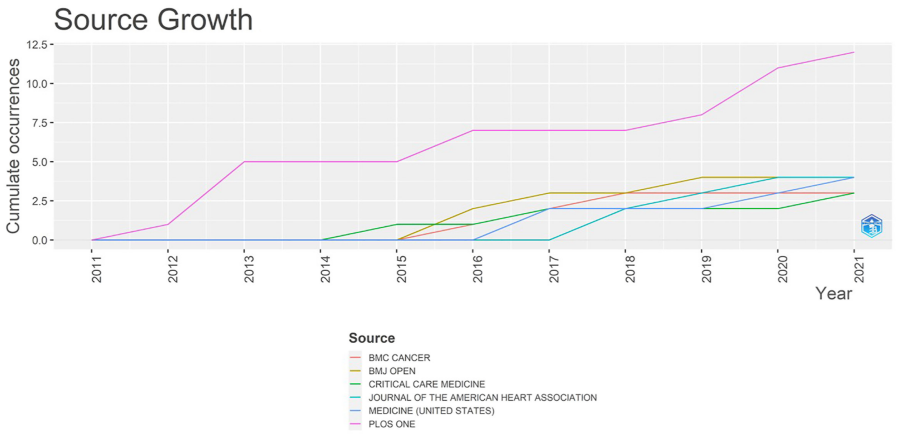
The most relevant authors who have published papers in this area have been demonstrated in Figure 6. The plot is created based on the number of documents published.

Figure 7 demonstrates the most relevant authors based on fractionalized frequency as a frequency measure. Fractional authorship quantifies an individual author's contribution to a published set of papers.

Element	H index	g_index	m_index
PLoS One	7	11	0.63
BMC Cancer	3	3	0.42
BMJ Open	3	3	0.42
Critical Care Medicine	3	3	0.37
Journal of the American Heart Association	3	4	0.6
Medicine (United States)	3	3	0.5
Annals of Thoracic Surgery	2	2	0.33
Cancer	2	2	0.16
Cancer Medicine	2	2	0.4
Clinical Cardiology	2	2	0.33

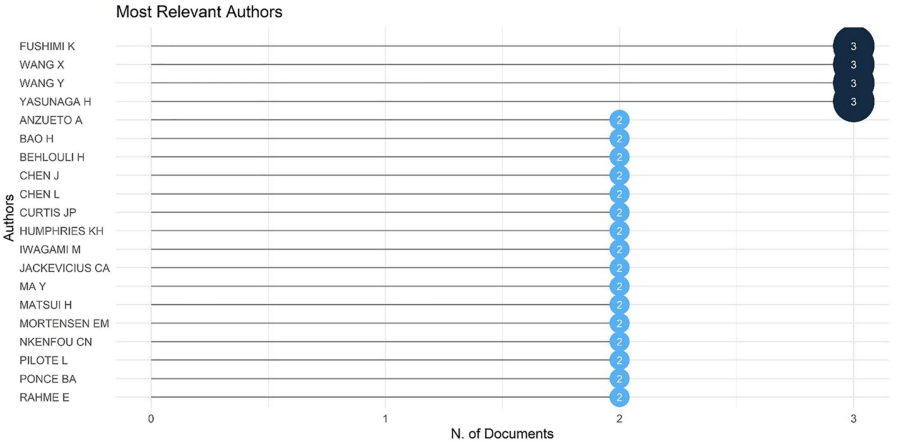
Source(s): Prepared by the Author (2023)

Table 3. Source local impact by H index, G index and M index



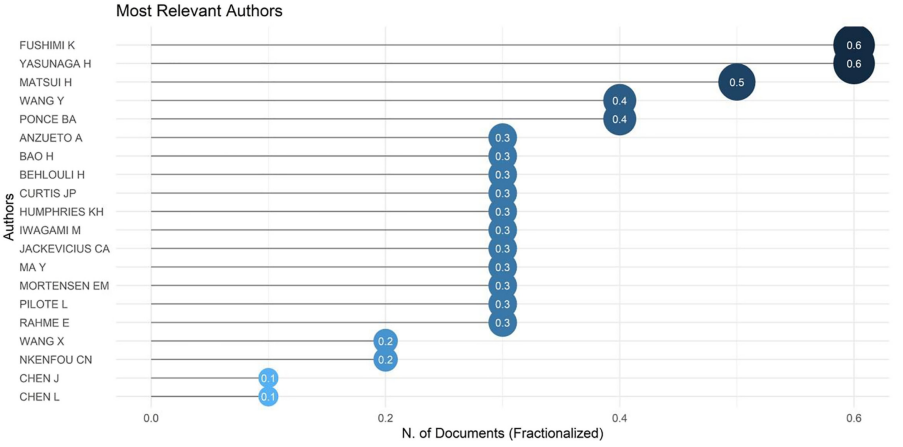
Source(s): Prepared by the author (2023)

Figure 5. Graph showcasing source growth between the years 2011 to 2021



Source(s): Prepared by the author (2023)

Figure 6.
A graph showcasing the most relevant authors based on the number of documents



Source(s): Prepared by the author (2023)

Figure 7.
Graph showcasing the most relevant authors based on the fractionalized frequency

Figure 8 lists the top 20 authors based on their local impact. The h index measure is used to obtain the authors' impact.

To represent the country analysis, Figures 9–11 present the analysis of the corresponding author country, the number of documents per country and number of citations per country.

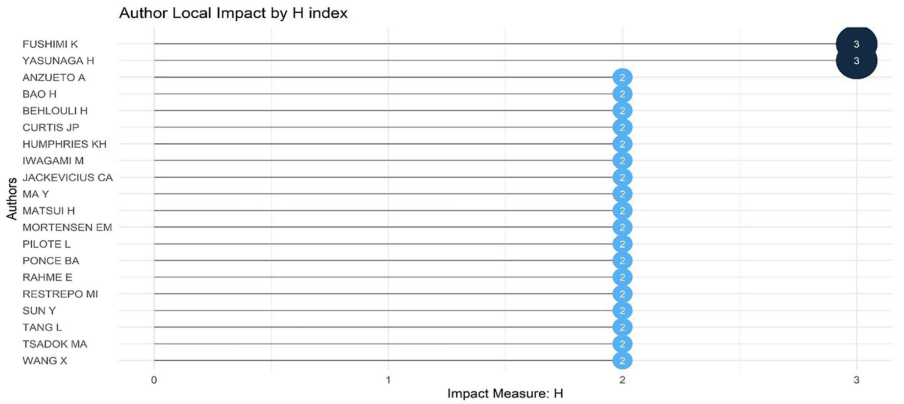
MCP: Multiple country publications is at least one coauthor from another country.

SCP: Single country publications is both the author and coauthors from the same country.

It is clear from the plot that, considering both SCP and MCP collaboration, the United States and China have the highest number of authors.

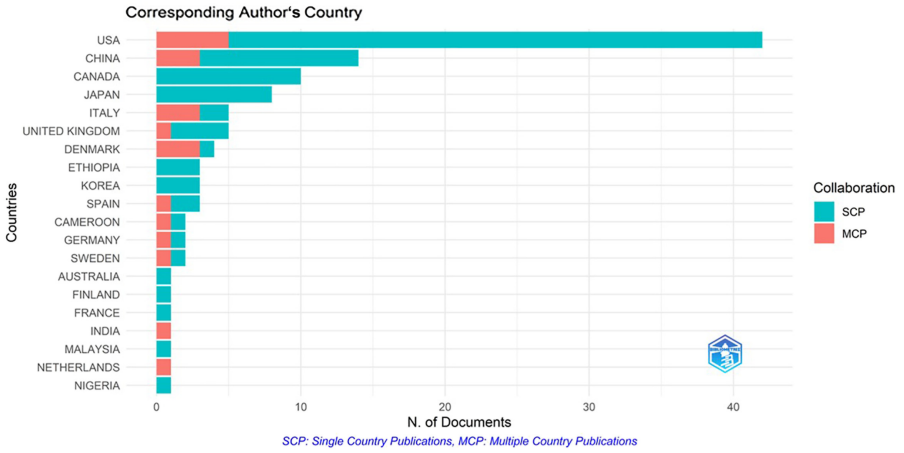
The figure below depicts a geographical map representing the country's scientific production. The color intensity is directly proportional to the number of publications. The darker the color, the higher the number of documents a country contributes.

The figure below shows the top 20 countries based on the number of citations obtained by a government. It is seen that the United States, Canada, Italy, China and the Netherlands are



Source(s): Prepared by the author (2023)

Figure 8.
Graph showcasing the top 20 authors based on local impact



Source(s): Prepared by the author (2023)

Figure 9.
Corresponding authors' country

the top five countries that have received the highest number of citations when compared to other countries.

4.3 Document analysis

The most globally cited articles in this area are depicted in Figure 12.

Figure 13 shows the word tree map. From the diagram, we can say that ML has most of its applications in mortality, HIV, cancer, diabetes, acute kidney disease, quality of care, depression and determining risk factors, as they have the highest percentage compared to other keywords.

The primary application of ML algorithms in healthcare management is to predict a patient's condition by studying the risk factors related to a disease. Apart from this, ML techniques are implemented to monitor one or more disorders (comorbidity) in a patient, which helps prevent the patient from the dreadful consequences of a disease. In addition, the

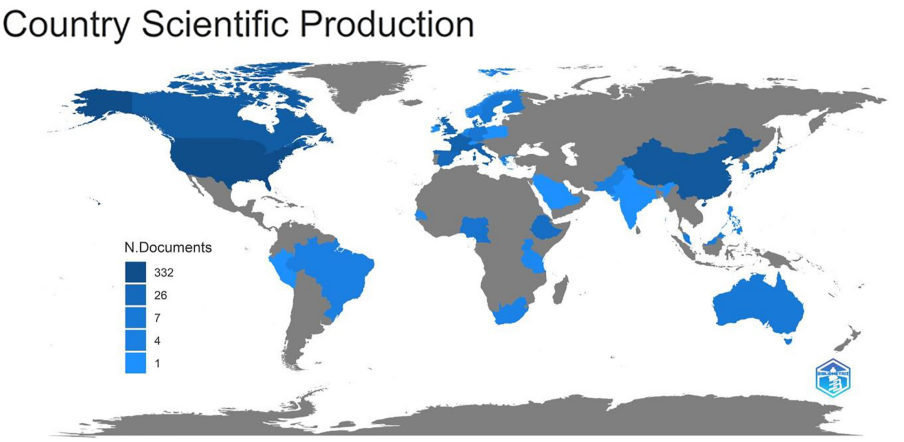


Figure 10.
Country scientific production

Source(s): Prepared by the author (2023)

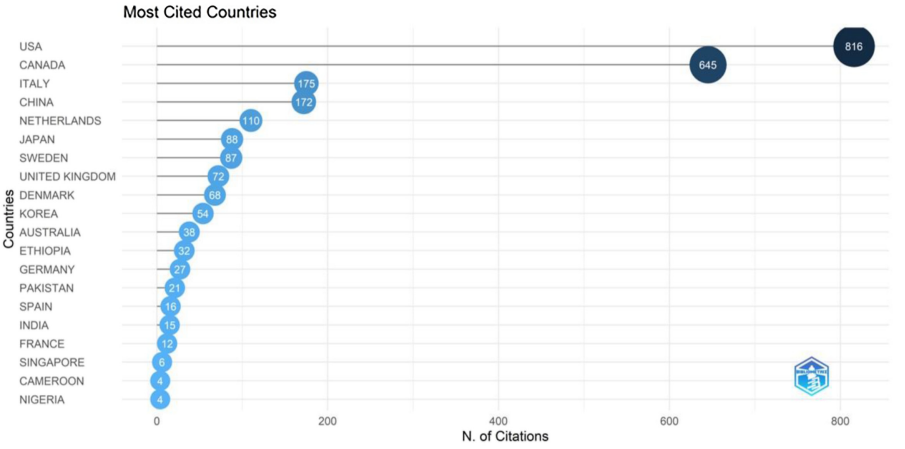


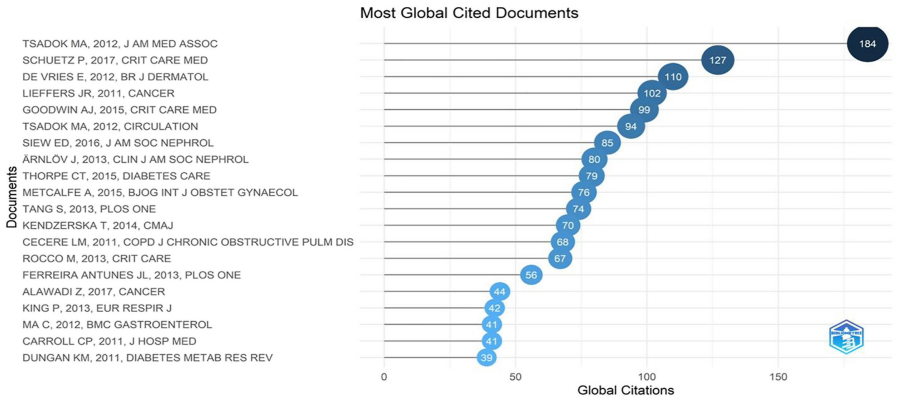
Figure 11.
Most cited countries

Source(s): Prepared by the author (2023)

seriousness of a patient (exacerbation) is determined by applying ML techniques such as fuzzy logic, which ultimately helps medical practitioners in decision-making. Management of beds and readmission of patients can be predicted by using time series analysis of ML, which helps increase hospital availability and management of resources.

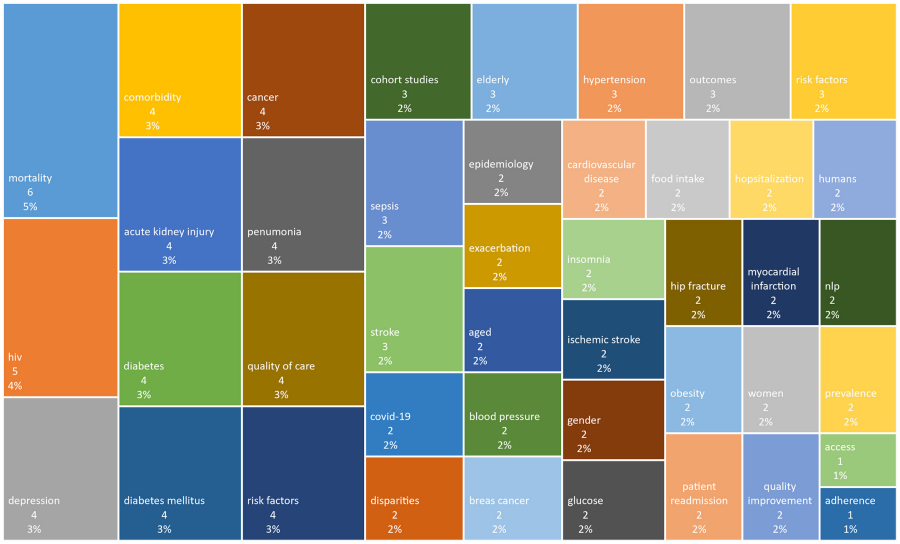
Figure 14 helps us understand the growth of prominent keywords over the years. This also helps determine the most trending topic in which ML is applied. It can be said that work in mortality and HIV disease has been increasing more over the years compared to other significant keywords.

Figures A4 and A5 in Appendix 4 depict the most frequently used keywords in the articles by the authors. Figures show keywords, namely mortality, HIV, cancer, diabetes mellitus, depression, hypertension, etc. The size of the keyword is directly proportional to the number of times a keyword is used. The larger the size, the more frequent a keyword is.



Source(s): Prepared by the author (2023)

Figure 12.
Most globally cited documents



Source(s): Prepared by the author (2023)

Figure 13.
Word tree map

A thematic map divided into four quadrants has been presented in Figure 15.

(1) Quadrant 1 (motor theme) represents high relevance and development degree.

(2) Quadrant 2 (niche theme) represents high relevance and low development degree.

(3) Quadrant 3 (emerging or declining theme) represents low relevance and development degree.

(4) Quadrant 4 (basic theme) represents low relevance and high development degree.

From quadrant 1, we can interpret that the topics, namely diabetes, cohort studies and hypertension, are the people who have done relevant and suitable work. Thus, to a greater extent, hospital management can rely on predictions and findings of ML algorithms in these areas.

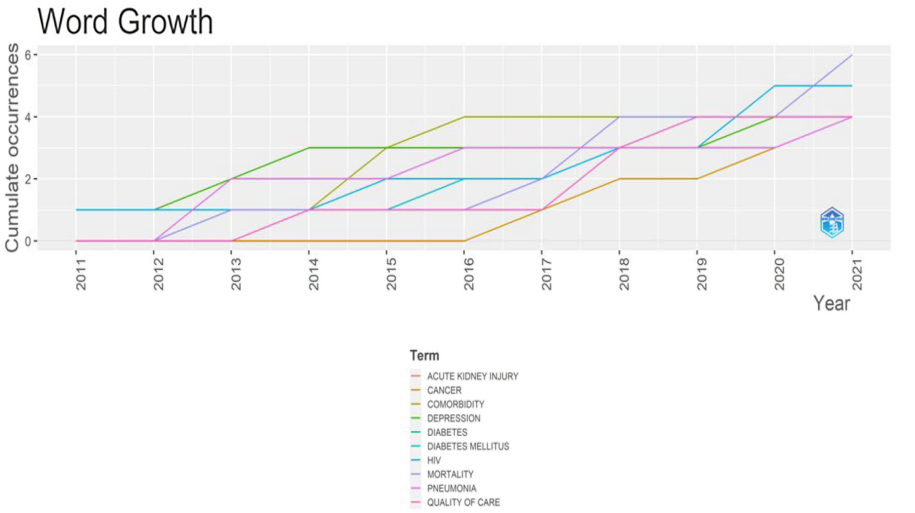


Figure 14.
Word growth

Source(s): Prepared by the author (2023)

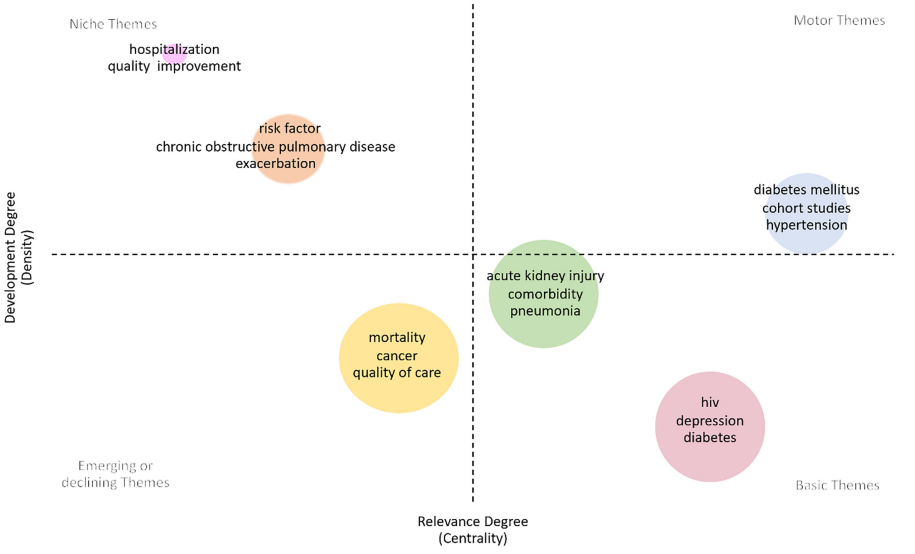


Figure 15.
Thematic map

Source(s): Prepared by the author (2023)

In quadrant 2 of the map, two clusters are found. Cluster 1, which includes "hospitalization" and "quality and quality improvement," tells us that though these topics are relevant to the study, very little work has been done in this area. As significantly less work has been done in these areas, upcoming ML researchers can take up these topics, whereas the second cluster consists of "risk factor," "chronic obstructive pulmonary disease" and "exacerbation" tells us that ML algorithms can be widely used to predict a particular patient's risk factors and seriousness. By applying ML algorithms, applications can be developed to improve the quality and efficiency of

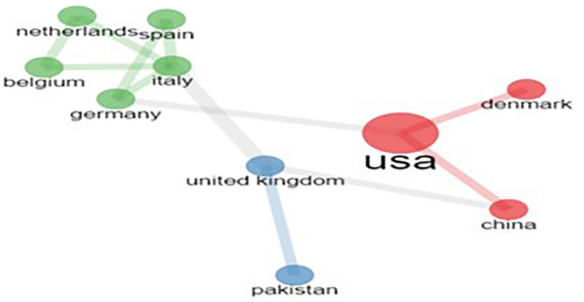
healthcare centers. Hospital management can develop telemedicine or remote monitoring applications for patient monitoring. There is enormous scope in this area.

Quadrant 3 reveals that much work has been done by applying ML in cancer, quality care and mortality. Hence, there needs to be more scope for development. Quadrant 4 reveals that though the topics in these quadrants, namely HIV, depression and diabetes, are less relevant, many papers are being published in these areas. The last cluster consists of topics such as acute kidney injury and pneumonia, which are moving toward motor themes with a higher scope for development.

4.4 Social structure analysis

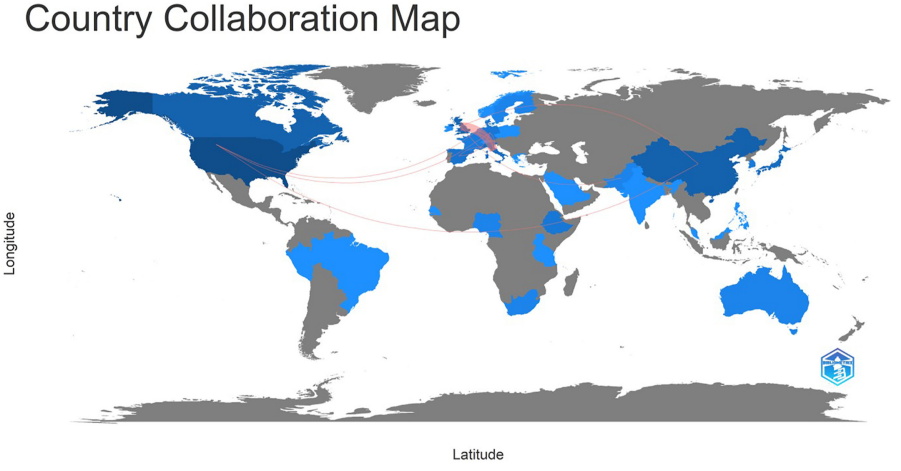
(1) Collaboration network (author)

The network (Figures 16 and 17) demonstrate the collaboration between authors from different countries contributing to this field. It is seen that three prominent clusters are formed. Countries that are together in a group means their country researchers mainly work together to prepare articles.



Source(s): Prepared by the author (2023)

Figure 16.
Graph showcasing collaboration network among the authors from different countries



Source(s): Prepared by the author (2023)

Figure 17.
Heat map showcasing country collaboration map

The first cluster consists of the United States, China and Denmark; the second consists of the United Kingdom and Pakistan; and the third comprises Germany, Italy, Spain, the Netherlands and Belgium.

(2) Country collaboration map

The graphs depict the links where country collaboration is shown. Table A2 in Appendix 1 shows the top 25 links with frequency links from one country to another.

5. Discussion

This systematic literature review examined the landscape of ML applications in healthcare strategic management. Our research question aimed to provide an overall analysis of healthcare institutions' adoption of ML and its implications. Our findings reveal significant growth in the integration of ML in the healthcare sector, as evidenced by the general annual scientific production. The main findings from our study have been synthesized in Table 4.

While existing related studies address specific applications or domains within healthcare, our study offers a comprehensive perspective on managing healthcare institutions. It doesn't limit itself to a particular medical condition or aspect but explores the broader spectrum of ML in healthcare management. This study provides a structured and data-driven approach to understanding the trends, influential authors and sources in the field, offering unique insights into the evolution of ML in healthcare management. In summary, this review shows a holistic approach to ML in healthcare management. It not only provides a comprehensive understanding of the field but also offers practical insights and addresses current challenges, making it a valuable addition to the existing body of research in the domain.

6. Conclusion and limitations

This review has shed light on the remarkable growth of ML in healthcare institutions. Through bibliometric analysis and a comprehensive literature review, emerging trends, influential authors and key critical applications were identified. This research provides insights for future developers and researchers to target areas with ML potential. Additionally, this paper highlighted managerial implications, showcasing how ML can enhance patient care, improve resource allocation, optimize emergency preparedness and promote innovative technologies like telemedicine and remote monitoring.

The theoretical contribution of the paper:

Findings	Description
Scientific production growth	Annual scientific production steadily increased from 2011 to 2021
Most relevant technologies	Wearable devices, mobile health app, telemedicine, blockchain, fog computing
Most relevant diseases	HIV, cancer, diabetes, acute kidney disease
Top journals for publishing most articles	*PLoS One* has twelve published articles followed by *BMJ Open*, *Journal of the American Association* and *Medicine (United States)*
Source impact by H index	*PLoS One* has the highest H index score of 7
Influential countries	The United States and China have the highest number of authors. United States, Canada, Italy, China and the Netherlands are the top five countries with the highest citations
Influential authors	Fushimi K, Wang X, Wang y, Yasunaga H, Anzueto A
Authors per document	7.52

Source(s): Prepared by the Author (2023)

Table 4. Main findings of our study

(1) The paper contributes by performing bibliometric analysis, revealing the substantial growth of ML in the healthcare sector. This analysis helps identify emerging areas and trends in scientific publications, categorizing topics that are essential and trending.

(2) It offers a comprehensive overview of implementing ML algorithms and techniques within healthcare institutions.

(3) The research identifies influential authors, sources, countries and keywords in the intersection of ML and healthcare management. This provides valuable insights into the key contributors and the most prominent themes in the field.

(4) Beyond analysis, the paper delves into the various applications of ML in healthcare, encompassing areas like IoT in healthcare, disease prediction, severity assessment, NLP, speech and language-based classification.

(5) The study reveals that journals such as *PLoS One*, *BMJ Open* and the *Journal of the American Heart Association and Medicine (United States)* are the leading publications in this field. It also highlights Fushimi K as the most prominent author based on fractionalized frequency and H index.

(6) By examining publication numbers, citations and document analysis, the paper establishes the United States as the leading country in ML and healthcare research. Furthermore, it identifies mortality and HIV as the top keywords, emphasizing the critical subjects addressed by authors in their published articles.

7. Managerial implications

(1) The study highlights the potential for ML to enhance medical support within healthcare institutions. It suggests that regression algorithms are particularly promising for this purpose.

(2) Hospital management can leverage time series ML algorithms to estimate the number of incoming patients, thus increasing hospital availability and optimizing resource allocation. This includes the preparation and stocking of essential medicines and drugs based on patient and disease trend analysis, reducing wastage.

(3) In response to the rise of telemedicine and remote monitoring, the paper recommends promoting these technologies in healthcare institutions. ML has been instrumental in the development of these systems. By embracing telemedicine and remote monitoring, healthcare management can facilitate the creation of online patient surveillance and monitoring systems, allowing for early medical intervention and ultimately improving the efficiency and effectiveness of medical services.

However, despite the valuable insights provided by this study, it's important to recognize certain limitations that should be addressed in future research and practical applications. This study predominantly concentrates on ML applications within specific healthcare domains, with some contexts still needing to be explored. For instance, future research could delve into additional dimensions, including ML applications in pharmaceuticals, personalized treatment plans or medical imaging diagnostics. Our study needs to explore various ML algorithm modeling techniques and their progressions and constraints within healthcare management. For instance, future research could be of analyzed metaheuristic algorithms for optimizing healthcare support. In conclusion, this study contributes to a better understanding the evolving landscape of ML applications in healthcare management. It offers valuable insights for future research and the improvement of healthcare systems.

References

Bolhasani, H., Mohseni, M., & Rahmani, A. M. (2021). Deep learning applications for IoT in health care: A systematic review. *Informatics in Medicine Unlocked*, *23*, 100550. doi: 10.1016/j.imu.2021.100550.

Born, J., Beymer, D., Rajan, D., Coy, A., Mukherjee, V. V., Manica, M., … Rosen-Zvi, M. (2021). On the role of artificial intelligence in medical imaging of COVID-19. *Patterns*, *2*(6), 100269. doi: 10.1016/j.patter.2021.100269.

Casey, A., Davidson, E., Poon, M., Dong, H., Duma, D., Grivas, A., … Alex, B. (2021). A systematic review of natural language processing applied to radiology reports. *BMC Medical Informatics and Decision Making*, *21*(1), 1–18. doi: 10.1186/s12911-021-01533-7.

Castelyn, G., Laranjo, L., Schreier, G., & Gallego, B. (2021). Predictive performance and impact of algorithms in remote monitoring of chronic conditions: A systematic review and meta-analysis. *International Journal of Medical Informatics*, *156*, 104620. doi: 10.1016/j.ijmedinf.2021.104620.

Chhetri, B., Goyal, L. M., & Mittal, M. (2023). How machine learning is used to study addiction in digital healthcare: A systematic review. *International Journal of Information Management Data Insights*, *3*(2), 100175.

Coccia, M. (2018). General properties of the evolution of research fields: A scientometric study of the human microbiome, evolutionary robotics, and astrobiology. *Scientometrics*, *117*(2), 1265–1283. doi: 10.1007/s11192-018-2902-8.

Coccia, M. (2020a). The evolution of scientific disciplines in applied sciences: Dynamics and empirical properties of experimental physics. *Scientometrics*, *124*, 451–487. doi: 10.1007/s11192-020-03464-y.

Coccia, M. (2020b). Deep learning technology for improving cancer care in society: New directions in cancer imaging driven by artificial intelligence, *Technology in Society*, (Vol. 60, pp. 1–11). February. art. n. 101198. doi: 10.1016/j.techsoc.2019.101198.

Coccia, M. (2020c). Destructive technologies for industrial and corporate change. In Farazmand, A. (Ed.), *Global Encyclopedia of Public Administration, Public Policy, and Governance*. Cham: Springer. doi: 10.1007/978-3-319-31816-5_3972-1.

Di Tanna, G. L., Wirtz, H., Burrows, K. L., & Globe, G. (2020). Evaluating risk prediction models for adults with heart failure: A systematic literature review. *PLoS One*, *15*(1), 1–23. doi: 10.1371/journal.pone.0224135.

Dugan, J., & Spetz, J. (2022). Containing healthcare costs. In *Health policy and advanced practice nursing: impact and implications* (p. 319).

Fernandes, R. C., Paredes, J. S., & Salinet, J. (2020). Detection and classification of cardiac arrhythmias by machine learning: A systematic review. *Computing in Cardiology*, 2020-Septe, 1–4. doi: 10.22489/CinC.2020.333.

Gad, I., Elmezain, M., Alwateer, M.M., Almaliki, M., Elmarhomy, G., & Atlam, E. (2023). Breast cancer diagnosis using a machine learning model and swarm intelligence approach. *2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC)*, Jeddah, Saudi Arabia (pp. 1–5), 2023, doi: 10.1109/ICAISC56366.2023.10085393.

Giannakopoulou, K., & Roussaki, I. (2022). Internet of Things Technologies and Machine Learning Methods for Parkinson's Disease Diagnosis, Monitoring and Management: A Systematic Review. *Sensors*, *22*(5), 1799.

Kumar, Y., Gupta, S., Singla, R., & Hu, Y. C. (2021). A systematic review of artificial intelligence techniques in cancer prediction and diagnosis. *Archives of Computational Methods in Engineering*, 1–28. doi: 10.1007/s11831-021-09648-w.

Kumar, Y., Koul, A., Singla, R., & Ijaz, M. F. (2022). Artificial intelligence in disease diagnosis: A systematic literature review, synthesizing framework, and future research agenda. *Journal of Ambient Intelligence and Humanized Computing*, 1–28. doi: 10.1007/s12652-021-03612-z.

Mohammadpour, S., Rabiei, R., Shahbahrami, E., Fathisalari, K., Khakzad, M., & Langarizade, M. (2022). The effects of implementing clinical decision support systems on the diagnosis, treatment, and management of cancers: A systematic review. *Journal of Payavard Salamat*, *16*(2), 113–122.

Mueller, N., Anderle, R., Brachowicz, N., Graziadei, H., Lloyd, S. J., de Sampaio Morais, G., … Rasella, D. (2023). Model choice for quantitative health impact assessment and modelling: An expert consultation and narrative literature review. *International Journal of Health Policy and Management*, *12*(Issue 1), 1–13, doi: 10.34172/ijhpm.2023.7103.

Oladimeji, O. O. (2023). Machine learning in smart health research: A bibliometric Analysis. *International Journal of information Science and management*, *21*(1), 117–126.

Ortíz-Barrios, M. A., Coba-Blanco, D. M., Alfaro-Saíz, J. J., & Stand-González, D. (2021). Process improvement approaches for increasing the response of emergency departments against the COVID-19 pandemic: A systematic review. *International Journal of Environmental Research and Public Health*, *18*(16). doi: 10.3390/ijerph18168814.

Pashazadeh, A., & Navimipour, N. J. (2018). Big data handling mechanisms in the healthcare applications: A comprehensive and systematic literature review. *Journal of Biomedical Informatics*, *82*, 47–62. doi: 10.1016/j.jbi.2018.03.014.

Pervez, M. S., Dutta, P., & Parihar, T. S. (2022). Digitalisation of health communication: History and current scenario. *Specialusis Ugdymas*, *1*(43), 3944–3966.

Raj, S. N. V., & Manivannan, S. K. (2022). Machine learning models to predict covid-19 vaccination intention: An Indian study. *International Journal of Professional Business Review*, *7*(6), e0977. doi: 10.26668/businessreview/2022.v7i6.e977.

Saleem, F., Al-ghamdi, A. S. A., Alassafi, M. O., & Alghamdi, S. A. (2022). Machine learning, deep learning, and mathematical models to analyze forecasting and epidemiology of COVID-19: A systematic literature review. *International Journal of Environmental Research and Public Health*. *19*(9), 5099.

Saxena, S., Deogaonkar, A., Pais, R., & Pais, R. (2023). Workplace productivity through employee sentiment analysis using machine learning. *International Journal of Professional Business Review*, *8*(4), e01216. doi: 10.26668/businessreview/2023.v8i4.1216.

Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., & Biancone, P. (2021). The role of artificial intelligence in healthcare: A structured literature review. *BMC Medical Informatics and Decision Making*, *21*(1), 1–23. doi: 10.1186/s12911-021-01488-9.

Shams, K., & Kadow, A. (2023). Subjective well-being, health and socio-demographic factors related to COVID-19 vaccination: A repeated cross-sectional sample survey study from 2021–2022 in urban Pakistan. *International Journal of Environmental Research and Public Health*, *20*(16), 6545.

Vigo, I., Coelho, L., & Reis, S. (2022). Speech-and Language-Based classification of Alzheimer's disease: A systematic review. *Bioengineering*, *9*(1). doi: 10.3390/bioengineering9010027.

Von Gerich, H., Moen, H., Block, L. J., Chu, C. H., DeForest, H., Hobensack, M., … Peltonen, L. M. (2022). Artificial intelligence -based technologies in nursing: A scoping literature review of the evidence. *International Journal of Nursing Studies*, *127*, 104153. doi: 10.1016/j.ijnurstu.2021.104153.

Xu, Q., Wang, L., & Sansgiry, S. S. (2020). A systematic literature review of predicting diabetic retinopathy, nephropathy, and neuropathy in patients with type 1 diabetes using machine learning. *Journal of Medical Artificial Intelligence*, *3*(March). doi: 10.21037/jmai.2019.10.04.

Zhang, T., Tan, T., Samperna, R., Li, Z., Gao, Y., Wang, X., … & Mann, R. M. (2023). Radiomics and artificial intelligence in breast imaging: a survey. *Artificial Intelligence Review*, 1–36. doi: 10.1007/s10462-023-10543-y.

Further reading

Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., & … Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. *Artificial Intelligence Review*, *56*, 1–97. doi: 10.1007/s10462- 023-10466-8.

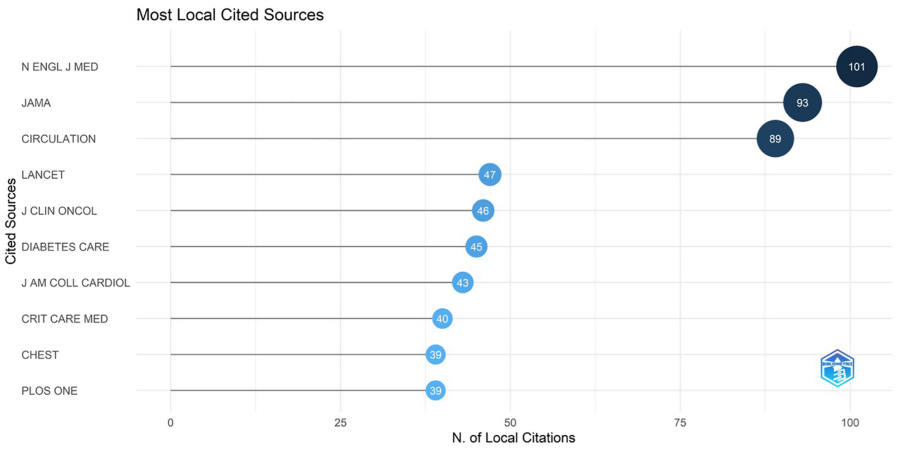
Appendix 1

Description	Results
Main information about the data	
Timespan	2011:2021
Sources (journals, books, etc.)	92
Documents	123
Average years from publication	5.28
Average citations per document	22.48
Average citations per year per doc	3.037
References	4,427
Document types	
Article	123
Document contents	
Keywords Plus (ID)	1,959
Author's Keywords (DE)	332
Authors	
Authors	925
Author appearances	955
Authors of single-authored documents	0
Authors of multiauthored documents	925
Authors collaboration	
Single-authored documents	0
Documents per author	0.133
Authors per document	7.52
Coauthors per documents	7.76
Collaboration index	7.52

Source(s): Prepared by the Author (2023)

Table A1. Important bibliometric statistics

Appendix 2
Most cited resources and source clustering



Source(s): Prepared by the author (2023)

Figure A1. Graph showcasing the most locally cited sources

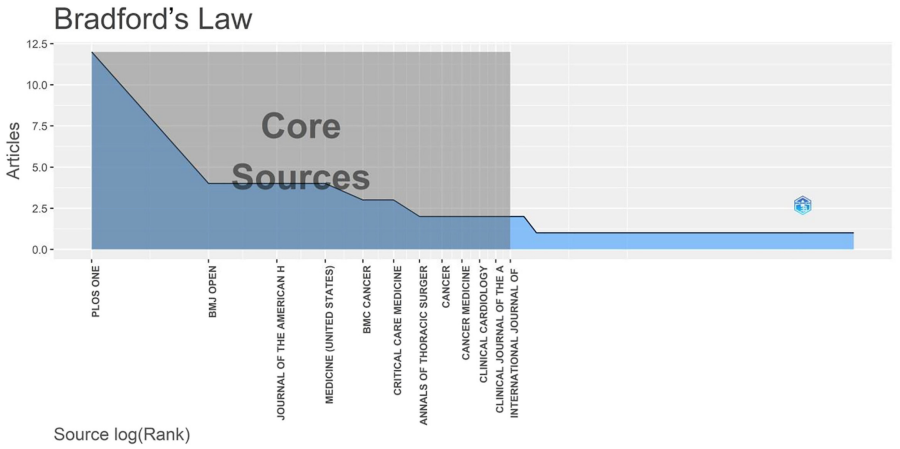


Figure A2. Graph showcasing source clustering

Source(s): Prepared by the author (2023)

Figure A2 represents the source clustering through Bradford's Law. Bradford's Law: "If the journals are arranged in descending order based on the number of articles, then a simple geometric series is formed by successive zones of periodicals containing the same number of articles on the subject." The center is called the nucleus, which shows the course sources. Out of 92 references, only 12 citations were found as core sources.

Appendix 3
H index

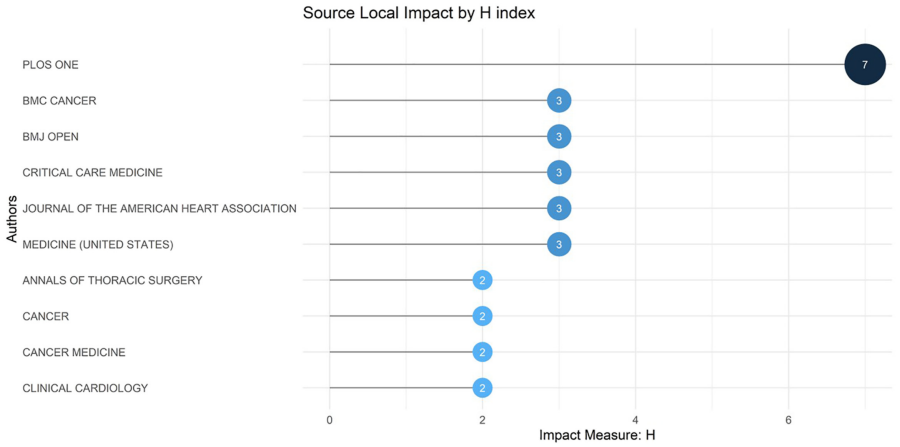


Figure A3. Graph showcasing source local impact by H index

Source(s): Prepared by the author (2023)

Figure A3 depicts the sources and their respective H index value. This shows that *PLoS One* has the highest H index value of 7.

Appendix 4
Keyword Analysis

Figure A4.
Word cloud

Source(s): Prepared by the author (2023)

The graph below represents the most relevant keywords used in the area.

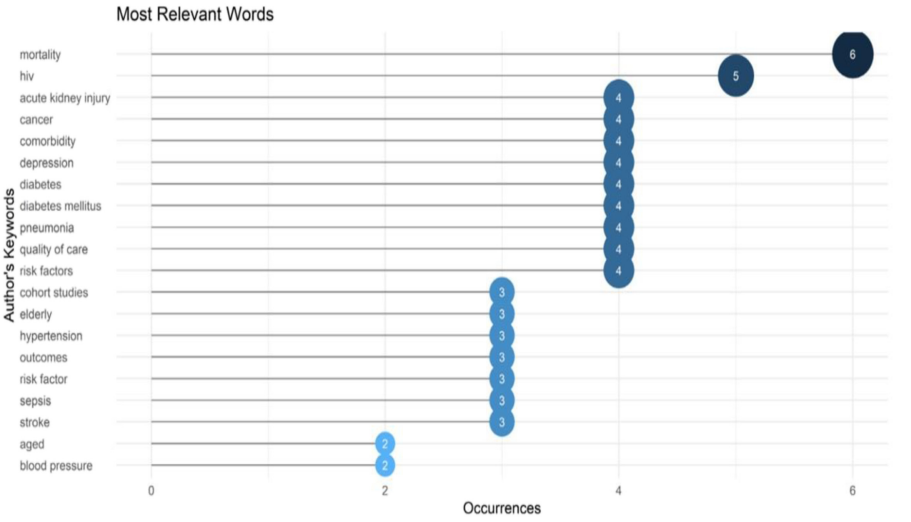


Figure A5.
Most frequent keywords

Source(s): Prepared by the author (2023)

From	To	Frequency
Italy	United Kingdom	3
Belgium	Netherlands	2
China	United Kingdom	2
Italy	Belgium	2
Italy	Germany	2
Italy	Netherlands	2
Italy	Spain	2
Spain	Germany	2
United Kingdom	Pakistan	2
United States	China	2
United States	Denmark	2
United States	Germany	2
Belgium	Greece	1
Belgium	Malta	1
Belgium	Poland	1
Belgium	Senegal	1
Cameroon	Belgium	1
Cameroon	Senegal	1
Canada	Czech Republic	1
Canada	France	1
Canada	Germany	1
Canada	Ireland	1
Canada	Italy	1
Canada	Pakistan	1
Canada	Singapore	1

Source(s): Prepared by the Author (2023)

Table A2. List of top 25 links between countries

It is shown that The United Kingdom has the highest level of collaboration with authors from different countries. Authors from the United Kingdom have collaborated with authors from Italy, China and Pakistan. Italy, the United States and Canada have collaboration and collaborated moderately with authors from various countries. Several countries, such as the Czech Republic, France and Ireland, have limited collaboration, with most of their collaborations resulting in only one instance. These countries are involved in fewer international research projects or joint publications.

Corresponding author
Sand Mohammad Salhout can be contacted at: ssalhout@staff.alquds.edu