ADBA
Computers and Electronics in Medicine
Cutting-Edge Scientific Solutions

e-ISSN: 3023-8609
RESEARCH ARTICLE
Vol.2 / No.1 / 2025 / pp.6-14
https://doi.org/10.69882/adba.cem.2025012

Bibliometric Analysis of Publications on Clinical Studies Leveraging Natural Language Processing During 2000 - 2023

Ömer Faruk Akmeşe[*,1] **and Buğra Bağcı**[β,2]
[*]Hitit University, Faculty of Engineering, Department of Computer Engineering, 19030, Corum, Turkiye, [β]Hitit University, Faculty of Engineering, Faculty of Economics and Administrative Sciences, 19030, Corum, Turkiye.

ABSTRACT The number of clinical studies using natural language processing is quite large. Therefore, it is important to examine in depth the development of clinical studies using Natural Language Processing over the years. However, there are a limited number of studies in the literature examining the research status of this field. The article presents a bibliometric analysis of studies on the keywords "clinical AND studies AND natural AND language AND processing" indexed in Scopus between 2000 and 2023. This study aims to evaluate academic outputs in the relevant field quantitatively, make sense of the data, reveal the state of scientific knowledge in the field, and give scientists a general perspective on the subject. Bibliometrix and Microsoft Excel programs were used for bibliometric analysis. Nineteen thousand two hundred seventy-three different authors identified a total of 4535 studies. 77.5% of these studies were research articles (3516), 14.8% were conference papers (669), 6.8% were reviews (307), and 0.9% were book chapters (43). Journal of Biomedical Informatics was the journal in which the most studies were published, with 226 articles. Only the United States (2637) contributed 58.1% to the studies. Liu, H. was the most prolific author, with 85 articles. Harvard Medical School was the most productive institution, with 304 studies. The most cited article was Discontinuation of Statins in Routine Care Settings, A cohort study.

KEYWORDS
Clinical studies
Natural language processing
Bibliometric analysis
Citation analysis
Network analysis

INTRODUCTION

Natural Language Processing (NLP) is a branch of artificial intelligence and linguistics that enables computers to understand expressions or words written in human languages (Khurana *et al.* 2023). In the 1950s, NLP initially focused on rule-based methods to enable computers to understand natural language. However, these insufficient methods have transformed over time with the developments in machine learning methods (Nadkarni *et al.* 2011). NLP studies have recently been included in various fields, such as machine translation, e-mail spam detection, information extraction, summarization, and medical question-answering (Khurana *et al.* 2023).

Most clinical information sources contain significant amounts of information. But most of this information comes in unstructured form (Meystre and Haug 2005). NLP is extremely important in transforming unstructured information into structured information, improving healthcare, and advancing medicine (Wang *et al.* 2017). NLP has applications in medical information processing and rich research achievements (Chen *et al.* 2018). NLP medical applications include numerous research topics, such as its use for mental health (Le Glaz *et al.* 2021; Corcoran and Cecchi 2020), extraction of structured information from radiology reports (Casey *et al.* 2021), coding clinical notes (Tavabi *et al.* 2022), and monitoring Alzheimer's disease (Garcia *et al.* 2020).

NLP has significant potential in clinical trials. This technology is expected to help increase the efficiency and effectiveness of medical research. NLP-supported medical research is rapidly increasing and becoming more attractive. However, there are a limited number of studies examining the research status of clinical studies on NLP. Therefore, it is essential to conduct an in-depth

Manuscript received: 13 December 2024,
Revised: 24 January 2025,
Accepted: 24 January 2025.

[1]ofarukakmese@hitit.edu.tr (**Corresponding author**).
[2]bugrabagci@hitit.edu.tr

analysis to understand the latest developments in this field. This study aims to examine the academic output of NLP in clinical studies.

Bibliometric analysis is a field of research that deals with numerical analysis of scientific literature, used to search and analyze comprehensive scientific data. This analysis may include research publication frequency, citations, authors, and topics. The state of the art in a field of current scientific knowledge can be mapped using bibliometrics. Bibliometrics is an important tool for analyzing the output of scientists, collaborations between universities, the effects of science funding on research and development performance, and educational productivity. Therefore, theoretical and practical tools are needed to measure experimental data. Bibliometric analysis has increased in popularity in recent years as the availability and accessibility of software such as Gephi, Leximancer, VOSviewer, and Bibliometrix and scientific databases such as Scopus and Web of Science have increased. Bibliometric analysis can measure research outputs, identify trends, and evaluate research performance (Akmese 2022; Falagas *et al.* 2006; Moral-Muñoz *et al.* 2020; Sengupta 1992; Donthu *et al.* 2021).

Bibliometric methods are now considered scientific expertise and have become integral to research evaluation methodology, especially in scientific and applied fields (Ellegaard and Wallin 2015). Bibliometric methods are often used to process data (Wallin 2005). These methods have greatly benefited from computerized data processing. Accordingly, there has been a great increase in the number of publications in this field in recent years. Increasing data volume and more widespread use of computers were effective in this increase (Ellegaard and Wallin 2015).

This study covers top journals, institutions, keyword features in the field, citation network analysis, and review of top articles, and offers the potential to illustrate historical and geographic trends. This study aims to make sense of the large number of data obtained, to quantitatively evaluate the academic outputs of relevant research, to provide scientists in this field with a general perspective on the subject, and to reveal the state of scientific knowledge.

This study can make various contributions to the field of research in question. It can provide domain experts with a comprehensive overview of the research topic. It can help better understand research outcomes. In addition, it can provide researchers with the most important information about potential authors, institutions, journals, and countries. It can help identify research trends or track the popularity and importance of topics. Moreover, it can increase researchers' awareness when deciding on topic selection. It can also help improve the quality and efficiency of research. Finally, it can explain how the topic has developed over time.

MATERIAL AND METHODS

The Scopus database was preferred to collect bibliometric information. All journals in the Scopus database are reviewed annually to maintain high-quality standards (Kokol *et al.* 2021). It has been determined that Scopus offers its users a more comprehensive journal profile than other databases and provides faster results from more articles in citation analysis.

All publications indexed in Scopus (access date: 18.12.2023) between 2000 and 2023 regarding clinical studies using natural language processing were analyzed using bibliometric methods. "clinical AND studies AND natural AND language AND processing" were used as search keywords. Documents were searched by article title, abstract, and keywords. Scopus codes used in the search are as follows: TITLE-ABS-KEY (clinical AND studies AND natural AND language AND processing) AND PUBYEAR > 1999 AND PUBYEAR < 2024 AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ch"))

This search method found all articles published between 2000 and 2023 in the studies' title, abstract, and keywords in the Scopus database. The number of studies may increase when 2023 is completed. Microsoft Excel and Bibliometrix (Aria and Cuccurullo 2017) were used for bibliometric network visualizations.

ANALYSIS

Literature Distribution

Four thousand five hundred thirty-five publications of different genres from 2000 to 2023 were evaluated. These publication types are articles (3516, 77.5%), conference proceedings (669, 14.8%), reviews (307, 6.8%), and book chapters (43, 0.9%).

As seen in Figure 1, clinical studies using Natural Language Processing, "Medicine" (3534, 45%), "Computer Science" (1113, 14%), "Health Professions" (584, 7%), "Engineering" (536, 7%), "Biochemistry, Genetics and Molecular Biology" (435, 6%), "Neuroscience" (275, 3%) and "Others" (1430, 18%). The total number of studies is more than 4535 because a study can be matched in more than one category.

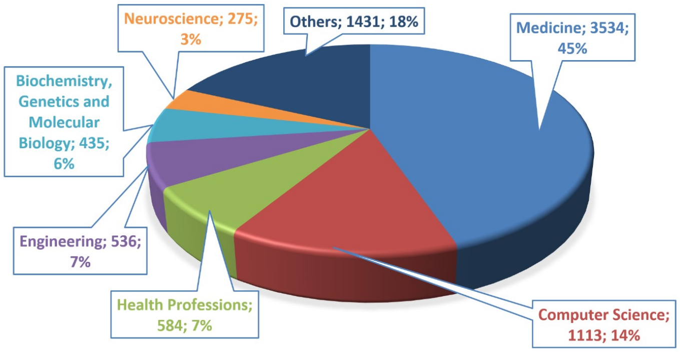

Figure 1 The distribution of subject areas

Development of Publications

The annual production graph of scientific studies of 4535 studies is shown in Figure 2. Despite some fluctuations, there has been an increase in the number of scientific studies in general. It is seen that the number of studies decreased in 2016. In the following years, the number of publications tends to increase continuously.

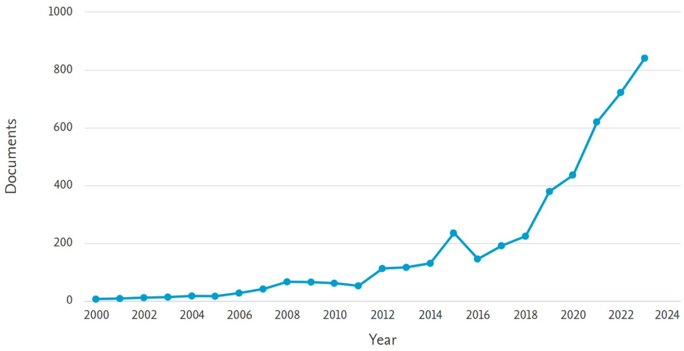

Figure 2 Document by year

Active Authors

Nineteen thousand two hundred seventy-three authors produced a total of 4535 works. The top five producing authors are Liu H. (85, 1.9%), Xu H. (74, 1.6%), Denny J.C. (47, 1%), Stewart R. (44, 1%), and Wu Y. (41, 0.9%). These authors were important research pioneers in their respective fields. Figure 3 shows the top 15 authors with the highest number of studies.

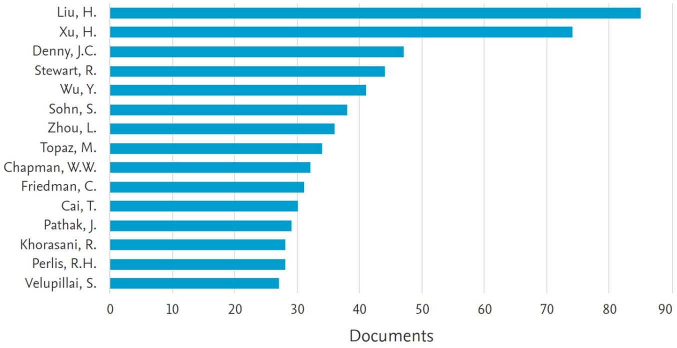

Figure 3 Top 15 authors with the highest number of studies

The collaboration network of the top 50 authors is shown in Figure 4. The size of the circles is directly proportional to the number of studies and collaboration between authors. Colors represent different clusters. The thickness of the lines expresses the strength of the collaboration between writers.

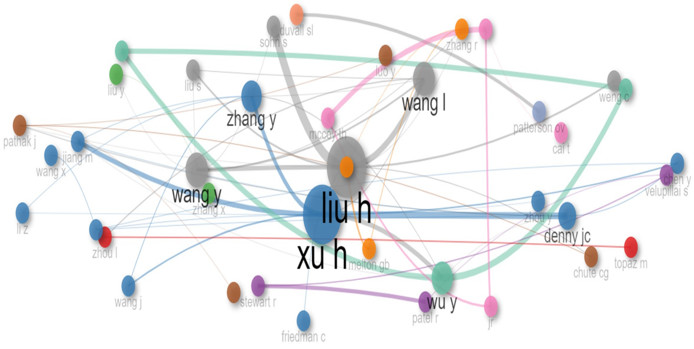

Figure 4 Authors Collaboration Network

Active Affiliation

The top 5 institutions that contributed the most to the literature were Harvard Medical School (304, 6.7%), Brigham and Women's Hospital (196, 4.3%), Massachusetts General Hospital (173, 3.8%), Mayo Clinic (169, 3.7%) and The University of Utah (137, 3%). Figure 5 shows the top 15 institutions that contributed the most according to the number of studies published by the institutions in the 2000-2023 period.

The collaboration network of the top 50 institutions is seen in Figure 6. The size of the circles is directly proportional to the number of studies and cooperation between institutions. Different colors represent clusters. The thickness of the lines expresses the strength of cooperation between institutions.

Active Journals

Table 1 shows the top 25 journals with the highest *h_index*. A total of 4535 studies were published in 1335 sources. 18.9% of the studies consist of the first five sources, and 38.5% comprise the 25 sources in Table 1.

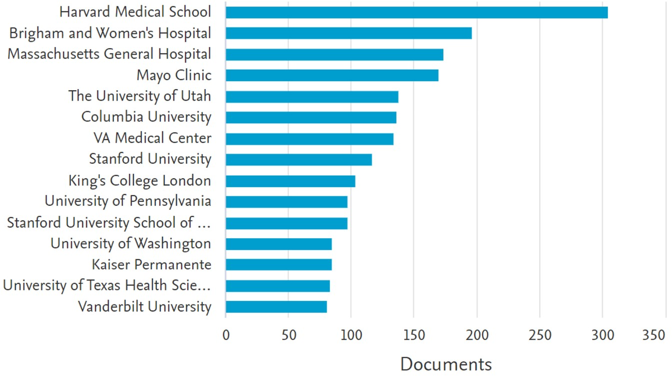

Figure 5 The top 15 organizations that contribute the most, according to the number of studies

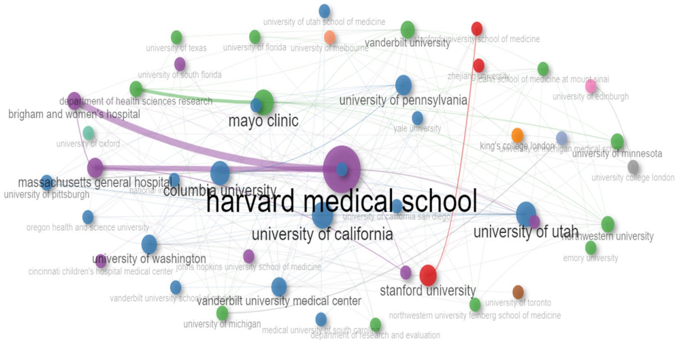
university of utah school of medicine
university of texas
university of florida
university of melbourne
vanderbilt university
university school of medicine
university of south florida
icahn school of medicine at mount sinai
department of health sciences research
university of pennsylvania
university of edinburgh
brigham and women's hospital
mayo clinic
yale university
university of oxford
king's college london
university of minnesota
university college london
massachusetts general hospital
columbia university
university of pittsburgh
harvard medical school
university of california san diego
oregon health and science university
university of california
university of utah
northwestern university
emory university
university of washington
cincinnati children's hospital medical center
johns hopkins university school of medicine
stanford university
vanderbilt university school
vanderbilt university medical center
university of toronto
northwestern university feinberg school of medicine
university of michigan
medical university of south carolina
department of research and evaluation

Figure 6 Institutions collaboration network

Figure 7 shows the increase in the number of publications of the top five journals according to their number of publications between 2000 and 2023. According to the chart, the Journal of Biomedical Informatics, Journal of the American Medical Informatics Association, Studies In Health Technology and Informatics, Journal of Medical Internet Research, and Jmir Medical Informatics sources are the most productive.

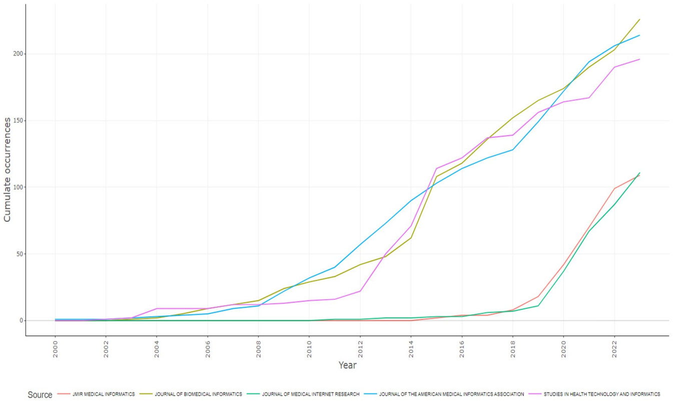

Figure 7 Top 5 journals with the highest number of articles

Table 1 Top 25 journals with h-index

No	Journal	h-index	g-index	m-index	TC	NP	PY_start
1	JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	53	90	2.208	9816	214	2000
2	JOURNAL OF BIOMEDICAL INFORMATICS	50	80	2.381	8374	226	2003
3	INTERNATIONAL JOURNAL OF RADIATION ONCOLOGY BIOLOGY PHYSICS	34	56	2.125	3228	65	2008
4	INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS	27	44	1.286	2279	97	2003
5	BMC MEDICAL INFORMATICS AND DECISION MAKING	26	43	1.3	2241	94	2004
6	AMIA ... ANNUAL SYMPOSIUM PROCEEDINGS / AMIA SYMPOSIUM	26	39	1.368	1889	90	2005
7	PLOS ONE	25	37	1.786	1668	94	2010
8	JOURNAL OF MEDICAL INTERNET RESEARCH	23	38	1.769	1785	111	2011
9	JMIR MEDICAL INFORMATICS	19	33	2.111	1383	109	2015
10	STUDIES IN HEALTH TECHNOLOGY AND INFORMATICS	16	29	0.727	1353	196	2002
11	RADIOLOGY	15	16	0.682	1656	16	2002
12	BMJ OPEN	13	32	1.444	1053	49	2015
13	AMIA ... ANNUAL SYMPOSIUM PROCEEDINGS. AMIA SYMPOSIUM	13	24	0.929	644	45	2010
14	ARTIFICIAL INTELLIGENCE IN MEDICINE	13	22	0.619	511	41	2003
15	JOURNAL OF THE AMERICAN COLLEGE OF RADIOLOGY	13	20	0.813	438	28	2008
16	JCO CLINICAL CANCER INFORMATICS	12	18	1.714	386	41	2017
17	METHODS OF INFORMATION IN MEDICINE	12	17	0.667	326	31	2006
18	COMPUTERS IN BIOLOGY AND MEDICINE	12	24	0.632	628	30	2005
19	COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE	11	14	0.733	221	18	2009
20	BMC BIOINFORMATICS	11	17	0.647	490	17	2007
21	JAMA NETWORK OPEN	10	18	1.667	371	32	2018
22	NPJ DIGITAL MEDICINE	10	19	1.667	454	19	2018
23	JAMIA OPEN	9	14	1.5	253	34	2018
24	APPLIED CLINICAL INFORMATICS	9	14	0.643	241	26	2010
25	IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS	9	18	1	324	21	2015
	TC: Total Citation, NP: Number of Publication, PY_start: Start of Publication Year						

Active Countries

Analysis showed that the articles covered 95 countries or territories. The publication numbers of the first 15 countries are shown in Figure 8. The United States ranked first with 2637 (58.4%) studies, considering the number of publications. China ranked second with 374 (8.3%) studies. The United Kingdom ranked third with 366 (8.1%) studies. Germany ranked 4th with 194 (4.3%), and Canada ranked 5th with 191 (4.2%).

The network visualization map for countries' international cooperation can be seen in Figure 9. The size of the circles is directly proportional to the number of studies and cooperation between countries. Colors represent different clusters. The thickness of the lines expresses the strength of cooperation between countries.

The geographical distribution of country collaboration for the overall study period is shown in Figure 10.

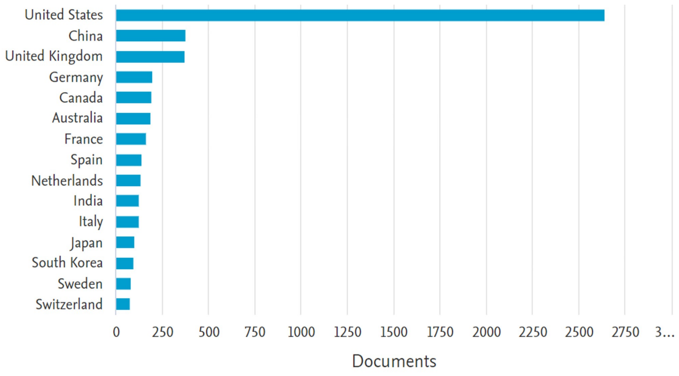

Figure 8 Bar chart showing the 15 most productive countries in the world

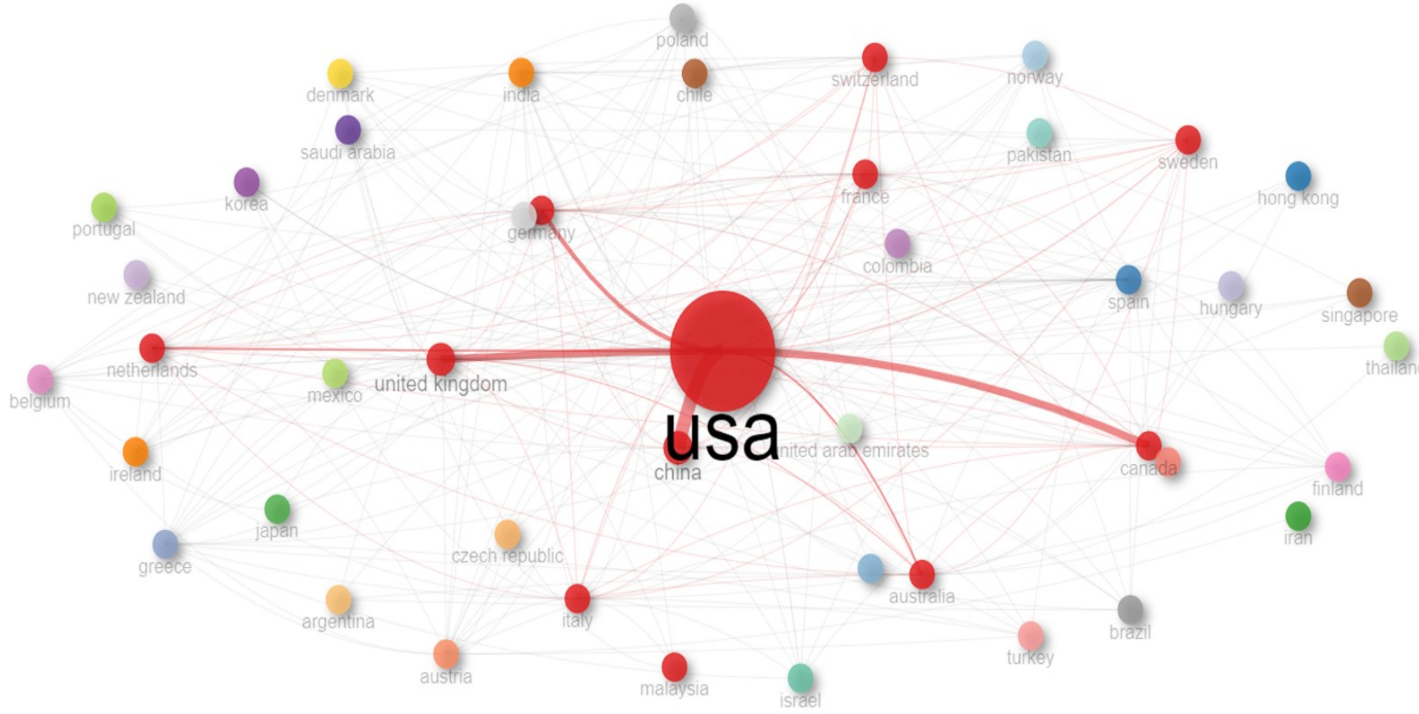

Figure 9 Network visualization map of countries' international cooperation

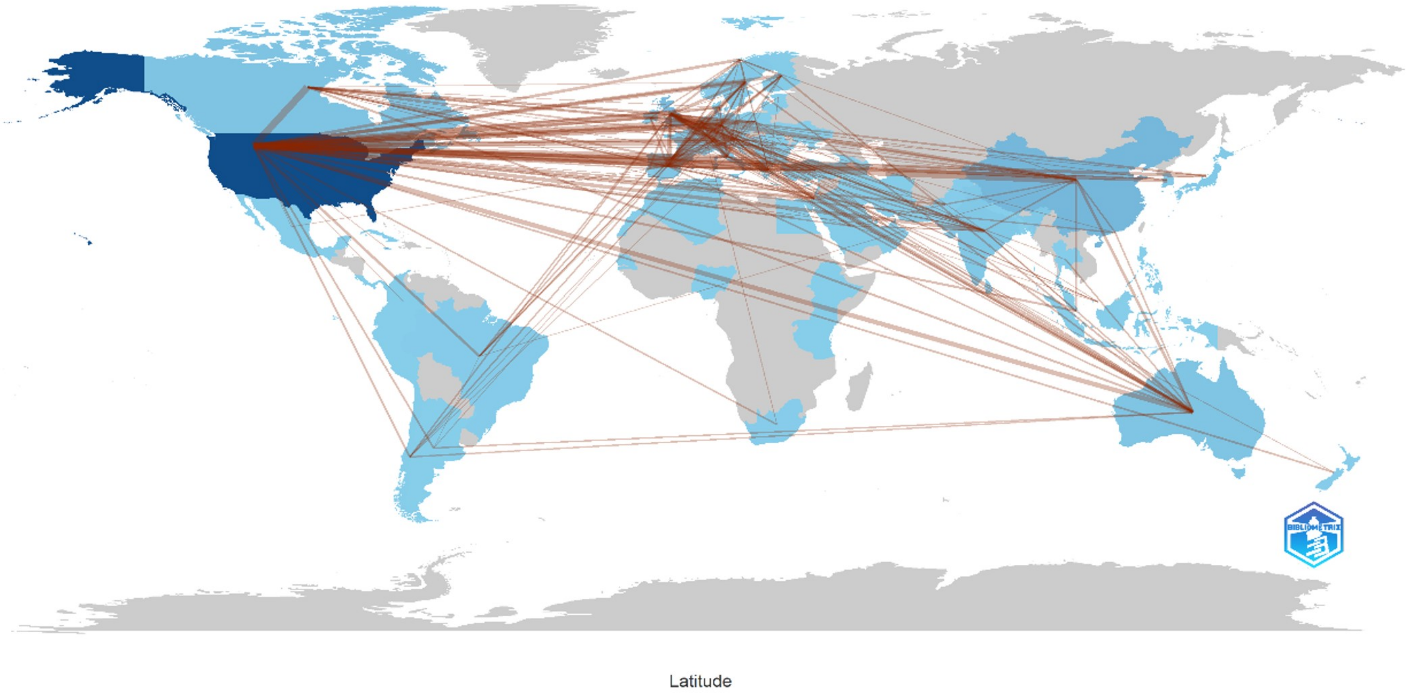

Figure 10 Country collaboration map

Citations

Citation status by publications is shown in Table 2. Figure 11 shows the co-citation network of the top 50 authors. As the size of the circle increases, the number of citations also increases. Colors represent different clusters. The thickness of the lines expresses the strength of the citation collaboration between authors.

Keyword Analysis

The 50 most used keywords in 4535 articles were visualized. The network visualization map of trend keywords obtained according to the topicality of publications is shown in Figure 12. As the size of the circle increases, the number of keyword uses also increases. The thickness of the lines expresses the strength of the connection

Table 2 Top 25 Papers by Total Citations (TC)

No	Paper	DOI	Total Citations	TC per Year	Normalized TC
1	ZHANG H, 2013, ANN INTERN MED	10.7326/0003-4819-158-7-201304020-00004	461	41.91	14.99
2	WANG Y, 2018, J BIOMED INFORMATICS-a	10.1016/j.jbi.2017.11.011	405	67.50	13.03
3	GOULD MK, 2015, AM J RESPIR CRIT CARE MED	10.1164/rccm.201505-0990OC	382	42.44	11.08
4	MILLER DD, 2018, AM J MED	10.1016/j.amjmed.2017.10.035	382	63.67	12.29
5	BEDI G, 2015, NPJ SCHIZOPHR	10.1038/npjschz.2015.30	380	42.22	11.02
6	LIANG H, 2019, NAT MED	10.1038/s41591-018-0335-9	349	69.80	13.51
7	MURFF HJ, 2011, J AM MED ASSOC	10.1001/jama.2011.1204	349	26.85	7.02
8	FRIEDMAN C, 2004, J AM MED INFORMATICS ASSOC	10.1197/jamia.M1552	343	17.15	7.25
9	BATES DW, 2003, J AM MED INFORMATICS ASSOC	10.1197/jamia.M1074	338	16.10	3.77
10	PONS E, 2016, RADIOLOGY	10.1148/radiol.16142770	337	42.13	9.61
11	PERERA G, 2016, BMJ OPEN	10.1136/bmjopen-2015-008721	323	40.38	9.21
12	SHIVADE C, 2014, J AM MED INFORMATICS ASSOC	10.1136/amiajnl-2013-001935	304	30.40	9.67
13	MEHTA N, 2018, INT J MED INFORMATICS	10.1016/j.ijmedinf.2018.03.013	284	47.33	9.13
14	HANAUER DA, 2015, J BIOMED INFORMATICS	10.1016/j.jbi.2015.05.003	282	31.33	8.18
15	CALVERT GA, 2003, J COGN NEUROSCI	10.1162/089892903321107828	280	13.33	3.12
16	SARKER A, 2015, J BIOMED INFORMATICS	10.1016/j.jbi.2014.11.002	273	30.33	7.92
17	TING DSW, 2019, PROG RETINAL EYE RES	10.1016/j.preteyeres.2019.04.003	273	54.60	10.57
18	TANG C, 2014, INT J RADIAT ONCOL BIOL PHYS	10.1016/j.ijrobp.2014.04.025	270	27.00	8.59
19	TITANO JJ, 2018, NAT MED	10.1038/s41591-018-0147-y	268	44.67	8.62
20	BOSSELER A, 2003, J AUTISM DEV DISORD	10.1023/B:JADD.0000006002.82367.4f	263	12.52	2.93
21	KISSLER J, 2006, PROG BRAIN RES	10.1016/S0079-6123(06)56008-X	260	14.44	4.62
22	NEAMATULLAH I, 2008, BMC MED INFORMATICS DECIS MAK	10.1186/1472-6947-8-32	258	16.13	7.23
23	RITCHIE MD, 2010, AM J HUM GENET	10.1016/j.ajhg.2010.03.003	250	17.86	5.02
24	HARKEMA H, 2009, J BIOMED INFORMATICS	10.1016/j.jbi.2009.05.002	247	16.47	5.34
25	KHO AN, 2011, SCI TRANSL MED	10.1126/scitranslmed.3001807	242	18.62	4.87

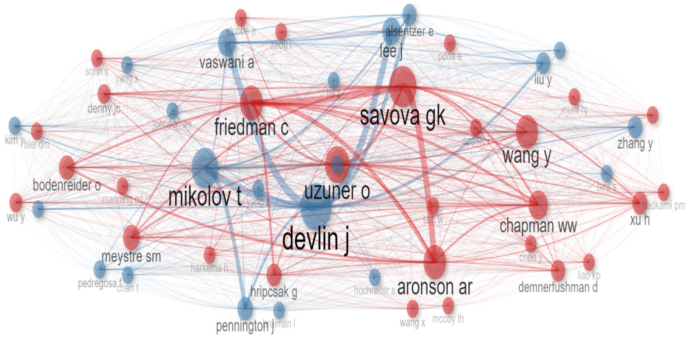

Figure 11 Author co-citation network

between keywords. Natural Language Processing, Human, Article, Humans, and Female were the articles' top 5 most frequently used keywords.

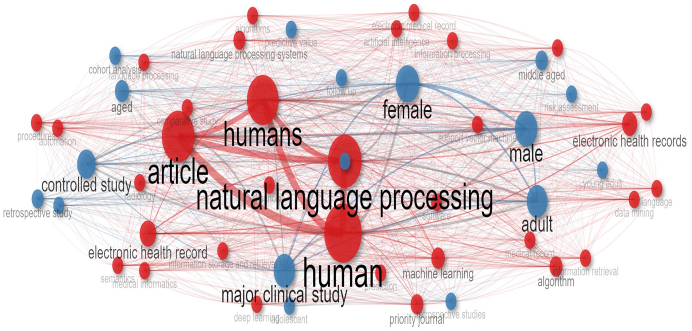

Figure 12 Keyword Analysis

Figure 13 shows the treemap of the most frequently repeated keywords and the number and percentage of repetitions of the 15 most used keywords.

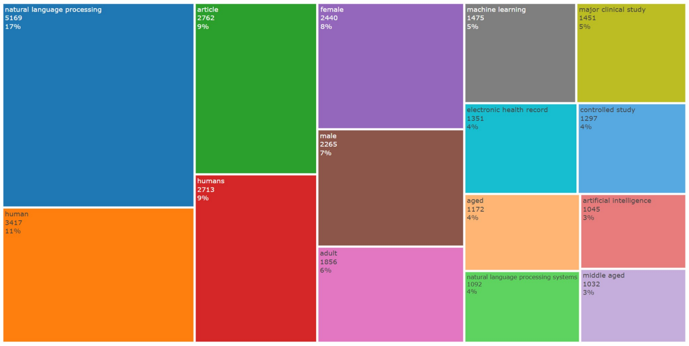

Figure 13 Treemap of most frequently repeated keywords

Thematic Evolution

The analysis of the evolution of keywords in the research is shown in Figure 14, showing the most frequently used keywords and their transformation over the years. The cut-off year was determined as 2017.

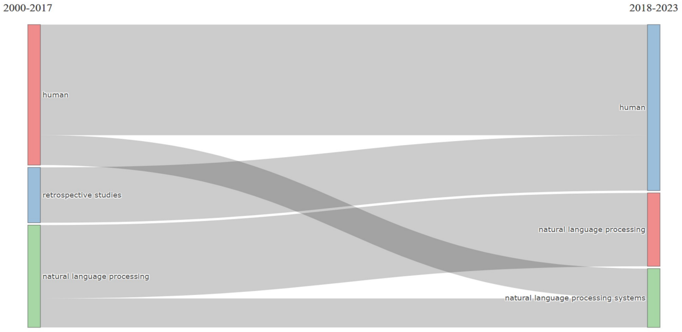

Figure 14 Thematic evolution by keywords

DISCUSSION

Although there was a general increase in the number of publications from 2000 to 2023, it is seen that the number of publications decreased significantly in 2016, and the increase continued in the following years. The first five subject areas of the studies are respectively *Medicine* (3534, 45%), *Computer Science* (1113, 14%), *Health Professions* (584, 7%), *Engineering* (536, 7%), *Biochemistry Genetics and Molecular Biology* (435, 6%) and *Neuroscience* (275, 3%).

The authors with the most publications on the subject are Liu H. (85, 1.9%), Xu H. (74, 1.6%), Denny J.C. (47, 1%), Stewart R. (44, 1%), and Wu Y. (41, 0.9%). The journals where the most published articles are the *Journal of Biomedical Informatics*, the *Journal of the American Medical Informatics Association*, *Studies in Health Technology and Informatics*, the *Journal of Medical Internet Research*, and *JMIR Medical Informatics*. 18.9% of the studies consist of the first five sources. The institutions that most contributed to the literature were *Harvard Medical School*, *Brigham and Women's Hospital*, *Massachusetts General Hospital*, *Mayo Clinic*, and *The University of Utah*. 21.6% of the studies consist of the top 5 organizations.

It is generally assumed that regional geographical location impacts collaboration when evaluating international collaborations. This assumption is that more cooperation can occur, especially between economically and scientifically developed countries. This assumption also appears valid for clinical studies using natural language processing. When the number of publications in a particular country is evaluated, it is seen that the countries with high economic power or large populations, such as the *USA*, *China*, the *United Kingdom*, *Germany*, and *Canada*, publish most studies on clinical studies using NLP. This situation aligns with the literature showing that academic productivity has a significant relationship with economic power (Demir 2019; Yıldırım and Demir 2019; Doğan and Kayır 2020). According to this literature, more academic studies are produced in these countries since more researchers and research resources exist in developed countries. Of course, it should be noted that regional geographical location is not the only factor that affects international cooperation. In addition, there may be factors such as cultural similarities, political relations, and shared interests.

According to the Scopus database, the first five most cited articles were (Zhang *et al.* 2013), *Annals of Internal Medicine* journal *"Discontinuation of statins in routine care settings: A cohort study"*; (Wang *et al.* 2018), *Journal of Biomedical Informatics* journal *"Clinical information extraction applications: A literature review"*; (Miller and Brown 2018), *American Journal of Medicine* journal *"Artificial Intelligence in Medical Practice: The Question to the Answer?"*; (Gould *et al.* 2015), *American Journal of Respiratory and Critical Care Medicine* journal *"Recent trends in the identification of incidental pulmonary nodules"*;

and (Bedi *et al.* 2015), *npj Schizophrenia* journal *"Automated analysis of free speech predicts psychosis onset in high-risk youths"*, respectively.

The articles' first five most frequently used keywords were *Natural Language Processing*, *Human*, *Article*, *Humans*, and *Female*. Limitations of the study: Although the Scopus database is advantageous compared to other databases regarding the number of publications, not all could be included. Additionally, since 2023 has not been completed, there may be a slight deficiency in the number of publications.

CONCLUSION

This study provides a holistic review of studies on clinical trials using Natural Language Processing between 2000-2023. According to the findings, it was observed that there was a decrease in the annual number of studies produced in 2016 and an increase in the following years. It was seen that the author with the most publications on the subject was Liu H., most articles were published in the JOURNAL OF BIOMEDICAL INFORMATICS, and the institution that contributed the most to the literature was Harvard Medical School. The most cited article (Zhang *et al.* 2013) was published in the Annals of Internal Medicine titled "Discontinuation of statins in routine care settings: a cohort study". The most productive countries in terms of the number of publications are developed or overpopulated countries. Participation of researchers in developing or underdeveloped countries in multinational studies may allow them to conduct further research on this subject.

Availability of data and material

Not applicable.

Conflicts of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethical standard

The authors have no relevant financial or non-financial interests to disclose.

LITERATURE CITED

Akmese, O. F., 2022 Bibliometric analysis of publications on chaos theory and applications during 1987 - 2021. Chaos Theory and Applications **4**: 169–178.

Aria, M. and C. Cuccurullo, 2017 bibliometrix: An r-tool for comprehensive science mapping analysis. Journal of Informetrics **11**: 959–975.

Bedi, G. *et al.*, 2015 Automated analysis of free speech predicts psychosis onset in high-risk youths. npj Schizophrenia **1**: 1–7.

Casey, A., E. Davidson, M. Poon, H. Dong, D. Duma, *et al.*, 2021 A systematic review of natural language processing applied to radiology reports. BMC Medical Informatics and Decision Making **21**: 1–18.

Chen, X., H. Xie, F. L. Wang, Z. Liu, J. Xu, *et al.*, 2018 A bibliometric analysis of natural language processing in medical research. BMC Medical Informatics and Decision Making **18**: 1–14.

Corcoran, C. M. and G. A. Cecchi, 2020 Using language processing and speech analysis for the identification of psychosis and other disorders. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging **5**: 770–779.

Demir, E., 2019 The evolution of spirituality, religion and health publications: Yesterday, today and tomorrow. Journal of Religion and Health **58**: 1–13.

Donthu, N., S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, 2021 How to conduct a bibliometric analysis: An overview and guidelines. Journal of Business Research **133**: 285–296.

Doğan, G. and S. Kayır, 2020 Global scientific outputs of brain death publications and evaluation according to the religions of countries. Journal of Religion and Health **59**: 96–112.

Ellegaard, O. and J. A. Wallin, 2015 The bibliometric analysis of scholarly production: How great is the impact? Scientometrics **105**: 1809–1831.

Falagas, M. E., A. I. Karavasiou, and I. A. Bliziotis, 2006 A bibliometric analysis of global trends of research productivity in tropical medicine. Acta Tropica **99**: 155–159.

Garcia, S. D. L. F., C. W. Ritchie, and S. Luz, 2020 Artificial intelligence, speech, and language processing approaches to monitoring alzheimer's disease: A systematic review. Journal of Alzheimer's Disease **78**: 1547–1574.

Gould, M. K., T. Tang, I. L. A. Liu, J. Lee, C. Zheng, *et al.*, 2015 Recent trends in the identification of incidental pulmonary nodules. American Journal of Respiratory and Critical Care Medicine **192**: 1208–1214.

Khurana, D., A. Koli, K. Khatter, and S. Singh, 2023 Natural language processing: state of the art, current trends and challenges. Multimedia Tools and Applications **82**: 3713–3744.

Kokol, P., H. B. Vošner, and J. Završnik, 2021 Application of bibliometrics in medicine: a historical bibliometrics analysis. Health Information and Libraries Journal **38**: 125–138.

Le Glaz, A., Y. Haralambous, D.-H. Kim-Dufor, P. Lenca, R. Billot, *et al.*, 2021 Machine learning and natural language processing in mental health: Systematic review. Journal of Medical Internet Research **23**: e15708.

Meystre, S. and P. J. Haug, 2005 Automation of a problem list using natural language processing. BMC Medical Informatics and Decision Making **5**: 1–14.

Miller, D. and E. Brown, 2018 Artificial intelligence in medical practice: the question to the answer? American Journal of Medicine **131**: 129–133.

Moral-Muñoz, J. A., E. Herrera-Viedma, A. Santisteban-Espejo, and M. J. Cobo, 2020 Software tools for conducting bibliometric analysis in science: An up-to-date review. Profesional de la Informacíón **29**.

Nadkarni, P. M., L. Ohno-Machado, and W. W. Chapman, 2011 Natural language processing: An introduction. Journal of the American Medical Informatics Association **18**: 544–551.

Sengupta, I. N., 1992 Bibliometrics, informetrics, scientometrics and librametrics: An overview. Libri **42**: 75–98.

Tavabi, N., M. Singh, J. Pruneski, and A. M. Kiapour, 2022 Systematic evaluation of common natural language processing techniques to codify clinical notes. medRxiv .

Wallin, J. A., 2005 Bibliometric methods: Pitfalls and possibilities. Basic and Clinical Pharmacology and Toxicology **97**: 261–275.

Wang, P., T. Hao, J. Yan, and L. Jin, 2017 Large-scale extraction of drug–disease pairs from the medical literature. Journal of the Association for Information Science and Technology **68**: 2649–2661.

Wang, Y., L. Wang, M. Rastegar-Mojarad, *et al.*, 2018 Clinical information extraction applications: a literature review. Journal of Biomedical Informatics .

Yıldırım, E. and E. Demir, 2019 Comparative bibliometric analysis of fertility preservation. Annals of Medical Research p. 1.

Zhang, H. *et al.*, 2013 Discontinuation of statins in routine care settings: A cohort study. Annals of Internal Medicine **158**: 526–534.

How to cite this article: Akmeşe, Ö. F., and Bağcı, B. Bibliometric Analysis of Publications on Clinical Studies Leveraging Natural Language Processing During 2000 - 2023. *Computers and Electronics in Medicine*, 2(1), 6-14, 2025.