Visualization Analysis of Current Situation and Trend of TCM Electronic Medical Records Research Based on CiteSpace

Gang Zhang, Shuanggui Tian(✉), and Ling Lin

Hubei University of Chinese Medicine, Wuhan, China
sunnytian610@hbtcm.edu.cn

Abstract. To explore the research status, hot spots and frontier trends in the field of electronic medical record of traditional Chinese medicine (TCM), and to provide reference for related researchers. Methods: The literatures related to TCM electronic medical records from January 1, 2014 to December 31, 2021 were retrieved from CNKI, Chongqing VIP, Wanfang and Web of Science databases. CiteSpace5.8 and other software were used to visually analyze the number of publications, publishing institutions, authors and keywords of the included literatures in each year. Results: A total of 431 valid Chinese literatures and 49 English literatures were included. The number of published literatures showed a steady upward trend. The key words of 0.1 are "data mining", "big data", "traditional Chinese medicine hall", "curative effect evaluation", etc., intermediary centrality of English literature & GT; 0.1 included "Traditional Chinese Medicine", "epidemiology", "diagnosis" and "cohort study". The analysis found seven clusters of Chinese keywords and nine clusters of English keywords. Conclusions: In recent years, the field of electronic medical records of traditional Chinese medicine at home and abroad has attracted increasing attention. The research hotspot has gradually shifted from the basic research of standardization of electronic medical records and the construction of information system to the optimization, design and application of data mining algorithm model based on electronic medical records of traditional Chinese medicine. The application research of extracting tacit knowledge, summarizing medication rules, and constructing intelligent auxiliary system of traditional Chinese medicine with functions of auxiliary clinical prediction, diagnosis, treatment, evaluation, etc.

Keywords: Electronic medical record of TCM · Visual analysis · CiteSpace · research hotspot

1 Introduction

TCM electronic medical record is the text, symbol, image and other data formed by digital technology in the process of TCM medical activities of medical staff. It is the concrete embodiment of the comprehensive application of TCM clinical principles, methods and prescriptions. It is also an important carrier to improve hospital medical service level and

D. Kumar et al. (Eds.): IEIT 2023, AHSSEH 10, pp. 57–68, 2023.
https://doi.org/10.2991/978-94-6463-230-9_8

promote the application and development of health medical big data. The research on electronic medical record of TCM is of great value to promote the development of TCM theory and clinical practice [1]. In this paper, the relevant literatures of TCM electronic medical records were retrieved, and the effective literatures were analyzed quantitatively and visually by CiteSpace software. At the same time, this paper draws a knowledge map based on the author and key words, so as to intuitively display research hotspots, with a view to more profoundly revealing the nature and laws of the research field of TCM electronic medical record [2]. The purpose of this study is to explore the research status, hot spots and frontier trends in the field of TCM electronic medical record, and provide reference for relevant researchers.

2 Data and Methods

In this paper, literature related to electronic medical records of TCM was retrieved from CNKI database (National Knowledge Infrastructure), Chongqing VIP database (China Science and Technology Journal Database), Wanfang database (Full text database of Chinese academic conference papers) and Web of Science database, and the publication period was limited to January 1, 2014 to December 31, 2021. In this study, CiteSpace 5.8 and other software are used to build a knowledge map and conduct visual analysis.

2.1 Data Sources

In this study, the Chinese literature is based on "traditional Chinese medicine" + "electronic medical record", and the English literature is based on "traditional Chinese medicine" and "electronic medical record". Some literature were excluded from the study, such as non authored literature, newspapers and other non-standard literature formats. According to the above search strategy and exclusion criteria, 313 literatures were retrieved from CNKI database, 265 literatures were retrieved from Wanfang database, and 81 literatures were retrieved from Chongqing VIP database. Then, the NoteExpress was used to de-duplicate the retrieved Chinese literature, and 431 valid Chinese literatures was obtained through manual reading and literature screening to supplement the title information. On the other hand, 49 English literatures were retrieved from the Web of Science database, and effective English literatures were selected by using Citespace to de-duplicate and manual reading of literature abstracts. Finally, 49 valid English literatures were obtained.

2.2 Research Methods

In this study, the bibliometrics method is adopted to analyze the collected literatures visually. Firstly, the collected literatures are converted to make them conform to the requirements of CiteSpace software. Secondly, the time parameter of the software is set to span from 2014 to 2021, and the year slice is 1 year. Then the authors and keywords of Chinese and English literature are analyzed. Finally, the corresponding knowledge map is generated. At the same time, Microsoft Excel 2016 software was used to make a statistical analysis of the annual number of publications, issuing institutions, keywords and other relevant information.

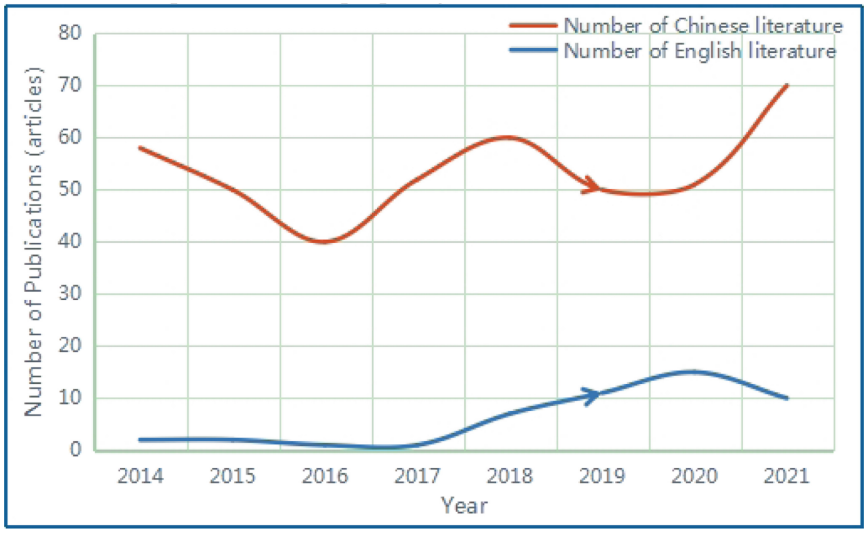

Fig. 1. Annual Number of literature publications in 2014–2021

3 Results

3.1 Analysis of the Number of Literature Published

The number of published literature is one of the important statistical indicators of bibliometrics. Figure 1 shows the number of Chinese and English literature publications in each year. As can be seen from the Fig. 1, that the number of published Chinese literature increased significantly in 2021, while the number of published English literature began to increase rapidly in 2017 and reached a peak in 2020. In general, the number of publications in the field of TCM electronic medical record is on the rise, indicating that scholars' attention to this field has remained stable. The rapid increase in the number of publications in 2021 May be related to the outbreak of novel coronavirus pneumonia and people's greater attention to health issues.

3.2 Author Analysis of the Literature

CiteSpace software was used to visualize the cooperation among authors, and two graphs were obtained. One is the cooperative diagram of Chinese authors, as shown in Fig. 2 (number of nodes 237, number of lines 537, network density 0.0192); the other is the cooperative diagram of English authors, as shown in Fig. 3 (number of nodes 135, number of lines 338, network density 0.0374). In the figure, the larger the authors' names are, the more papers they publish; the lines between authors represent the cooperative relationship between authors; the thicker the lines are, the more frequent the cooperation between authors [3].

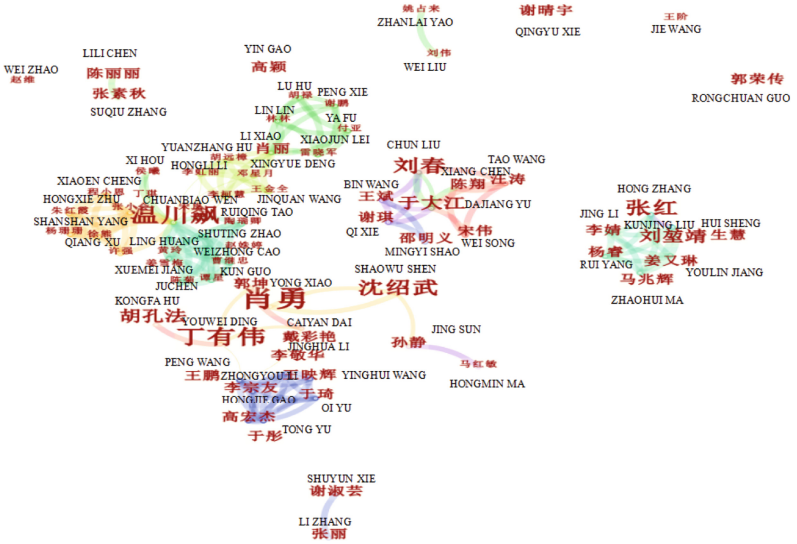

Fig. 2. Cooperative Network of Chinese Literature Authors

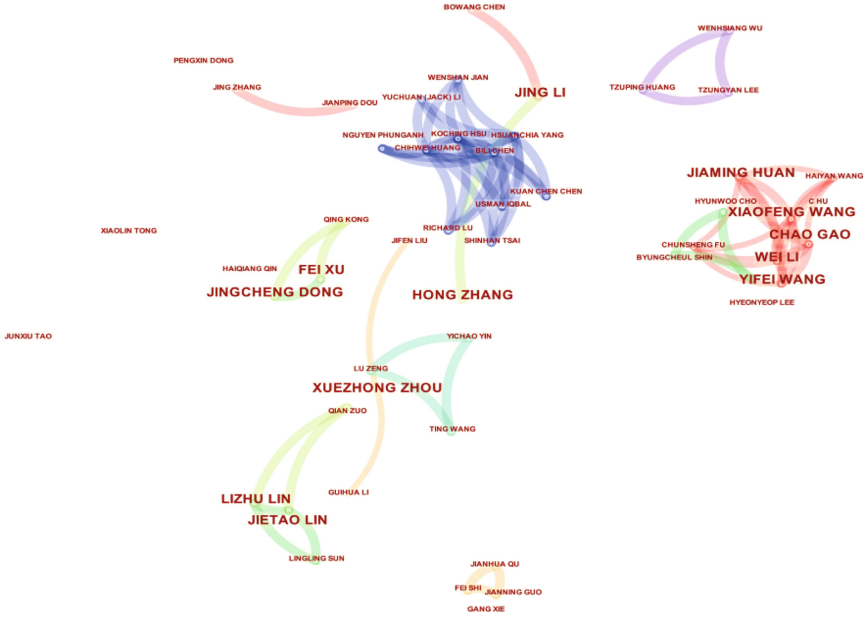

Fig. 3. Cooperative Network of English Literature Authors

3.3 Analysis of Author Units of Literature

The author made statistics on the number of publications of Chinese and English literature research institutions, and the information of the top five research institutions with

Table 1. Top five research institutions in terms of number of Literature issued

Chinese literature		English literature	
research institution	number of publications (articles)	research institution	number of publications (articles)
China Academy of Chinese Medical Sciences	45	China Academy of Chinese Medical Sciences	14
Hubei University of Chinese Medicine	29	Hubei University of Chinese Medicine	6
Chengdu University of Traditional Chinese Medicine	28	Beijing Jiaotong University	4
Nanjing University of Chinese Medicine	18	Chengdu University of Traditional Chinese Medicine	4
Beijing University of Chinese Medicine	15	Guangzhou University of Chinese Medicine	4

the number of publications is shown in Table 1. The top five research institutions with the number of Chinese literature publications were China Academy of Chinese Medical Sciences (45 articles), Hubei University of Chinese Medicine (29 articles), Chengdu University of Traditional Chinese Medicine (28 articles), Nanjing University of Chinese Medicine (18 articles) and Beijing University of Chinese Medicine (15 articles). The top five institutions with the number of articles published in English were China Academy of Chinese Medical Sciences (14 articles), Hubei University of Chinese Medicine (6 articles), Beijing Jiaotong University (4 articles), Chengdu University of Traditional Chinese Medicine (4 articles) and Guangzhou University of Chinese Medicine (4 articles). It can be seen that the China Academy of Chinese Medical Sciences occupies an important position in this research field.

3.4 Keyword Co-occurrence and Cluster Analysis

Keywords are the words that deeply summarize the subject content of the article and can reflect the core content of the literature. The co-occurrence analysis of keywords in the literature can explore the topic of the article and reflect the research status and hot spots in this field. Keyword cluster analysis is to cluster keywords with high similarity through clustering algorithm, and then give each keyword a value. The largest value in the same cluster is selected as the representative of this category, so as to visualize the basic knowledge structure in this field [4]. Keyword analysis was conducted on the retrieved Chinese and English literature, and keyword co-occurrence maps were obtained respectively, as shown in Fig. 4 and Fig. 6. In order to visually display the time span of literature in each cluster and the close relationship between clusters, LLR algorithm was used to cluster keywords. The time of publication of literature was taken as the X-axis

and the cluster number was taken as the Y-axis, and the clustering time line graphs of keywords in Chinese and English literature were drawn respectively, as shown in Fig. 5 and Fig. 7. The keyword network in Fig. 4 and Fig. 5 consists of 225 nodes and 340 lines. The network density is 0.0204, the modularization Q is 0.7915 (>0.3), and the average contour value S is 0.9769 (>0.7). The keyword network in Fig. 6 and Fig. 7 consists of 134 nodes and 385 lines. The network density is 0.0432, its modularization Q is 0.8268 (>0.3), and the average contour value S is 0.954 (>0.7). It is generally believed that the closer the module degree is to 1, the clearer the structure of the clustering results, and the closer the contour value is to 1, the higher the reliability of the clustering results. It can be seen that the clustering effect divided in Fig. 5 and Fig. 7 is better and the clustering results are credible. In the study, 7 Chinese keyword clusters and 9 English keyword clusters were also counted, and the clustering information was shown in Table 2 and Table 3.

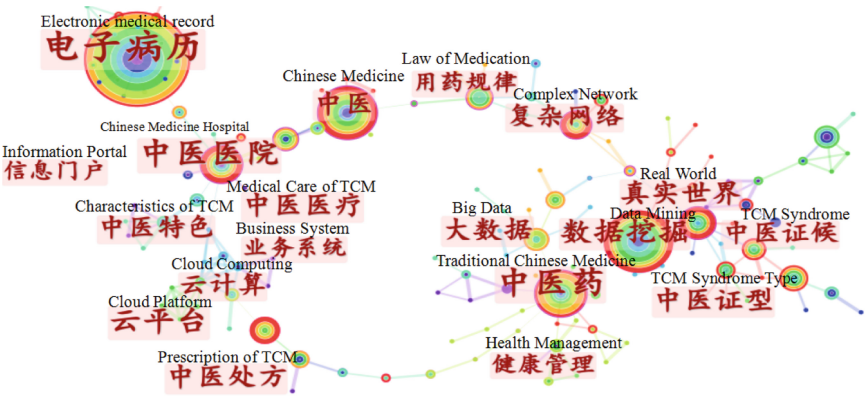

Fig. 4. Keyword co-occurrence map of Chinese literature

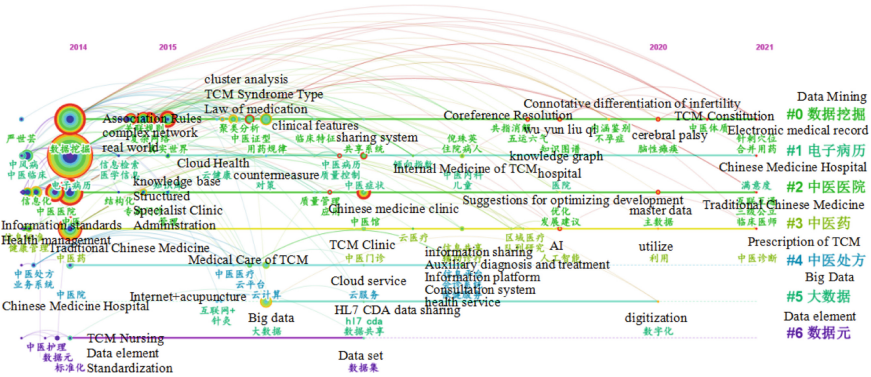

Fig. 5. Keyword clustering timeline map of Chinese literature

Fig. 6. Keyword co-occurrence map of English literature

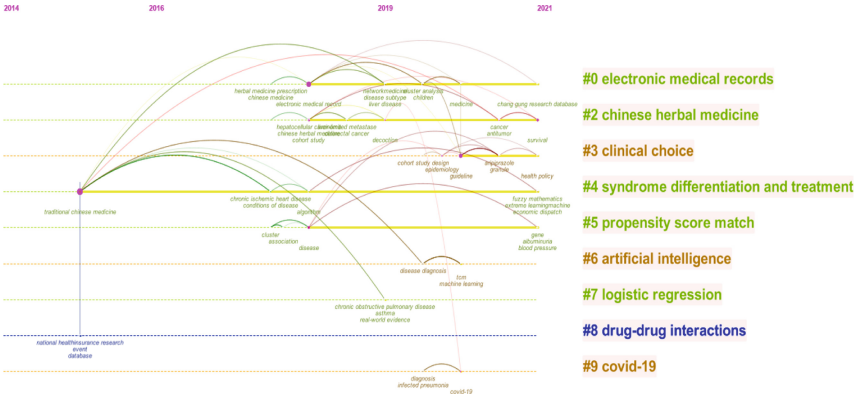

Fig. 7. Keyword clustering timeline map of English literature

3.5 Analysis of Keyword Emergence in Literature

Emergent keywords refer to the sudden high-frequency keywords in a certain stage, which can reflect the research status and the evolution of research hotspots in this field to a certain extent, so as to predict the development direction of research topics [5]. Based on Fig. 4 and Fig. 6, the sudden detection of keywords was carried out, and two keywords were obtained through the analysis of Chinese literature (see Fig. 8), indicating that since 2016, researchers in relevant fields have mostly taken electronic medical records of traditional Chinese medicine as the object and used data mining algorithms such as cluster analysis to mine the data rules of traditional Chinese medicine. However, the analysis of English literature shows that there are no emergent keyword

Table 2. Chinese keyword clustering information

Cluster No.	Cluster Label	Cluster Size	Time	Cluster Content
0	数据挖掘 data mining	40	2017	数据挖掘; 聚类分析; 用药规律; 真实世界; 关联规则 Data mining; Cluster analysis; Law of medication; real world; Association Rules
1	电子病历 Electronic medical record	31	2016	电子病历; 中医临床; 知识库; 决策支持; 中风病 Electronic medical record; TCM clinic; Knowledge base; Decision support; apoplexy
2	中医医院 Traditional Chinese Medicine Hospital	26	2016	中医医院; 中医; 临床路径; 管理; 中医特色 Chinese Medicine Hospital; chinese medicine; Clinical pathway; Administration; Characteristics of TCM
3	中医药 Traditional Chinese Medicine	19	2017	中医药; 人工智能; 健康管理; 队列研究; 区域医疗 Traditional Chinese Medicine; artificial intelligence; Health management; Cohort study; Regional medical treatment
4	中医处方 Prescription of Traditional Chinese Medicine	14	2015	中医处方; 名医医案; 云平台; 知识库系统; 服务总线 Prescription of TCM; Medical cases of famous doctors; Cloud platform; Knowledge base system; Service Bus

(*continued*)

Table 2. (*continued*)

Cluster No.	Cluster Label	Cluster Size	Time	Cluster Content
5	大数据 big data	8	2016	大数据; hl7 cda; 数据共享; 中医传承; apriori big data; hl7 cda; Data sharing; Inheritance of TCM; apriori
6	数据元 Data element	5	2014	数据元; 标准化; 中医护理; 信息; 电子病历 Data element; Standardization; TCM nursing; Information; Electronic medical record

nodes, which indicates that the activity degree of all academic branches in this field is relatively balanced. Sigma value is an index composed of intermediate centrality and burst value, which can be used to evaluate the innovation in this field. The more nodes with a high sigma value in a certain category, the higher the innovation in this category [6]. The sigma value in the keyword co-occurrence network in Fig. 6 is 1, indicating that the innovation in this field is not significant abroad.

4 Discuss

This study analyzed 431 Chinese and 49 English literatures on the topic of TCM electronic medical records from 2014 to 2021. The purpose of this study is to summarize and summarize the domestic and foreign research status, cooperation network and frontier hot spots in this field based on the annual number of publications, authors, institutions and keywords. Through the above analysis, it is found that the annual number of publications at home and abroad has an increasing trend, indicating that domestic and foreign researchers continue to deepen the research on electronic medical records of traditional Chinese medicine. However, the number of articles published in English is few, which indicates that Chinese researchers of TCM electronic medical records have not published enough in foreign languages.

Through the author's analysis, it is found that the characteristics of TCM electronic medical record research team are obvious. A research team with Shen Shaowu, Wang Yinghui, Wen Chuanbiao, Hu Kongfa, Zhang Hong and Liu Chun as the core has been formed in this field in China. The team of Shen Shaowu in Hubei University of Chinese Medicine focuses on the standardization and informationization of traditional Chinese medicine and emphasizes that electronic medical record of traditional Chinese medicine is the core of the informationization construction of traditional Chinese medicine hospitals under the new medical reform [7]. The authors of the English literature are mainly Chinese researchers, whose research directions mainly include the construction of electronic medical record system of traditional Chinese medicine, information extraction,

Table 3. English keyword clustering information

Cluster No.	Cluster Label	Cluster Size	Time	Cluster Content
0	electronic medical records	18	2019	electronic medical records; community detection; complex network; symptoms; disease subtypes
2	Chinese herbal medicine	14	2019	Chinese herbal medicine; overall survival; hepatocellular carcinoma; colorectal cancer; liver-limited metastases
3	clinical choice	14	2020	clinical choice; pharmacological intervention; epidemiology; health policy; neurology
4	syndrome differentiation and treatment	13	2018	syndrome differentiation and treatment; fuzzy mathematics; chronic ischemic heart disease; primary liver cancer; algorithm
5	propensity score match	12	2018	propensity score match; kidney; overexpression; performance status; prevention
6	artificial intelligence	9	2020	artificial intelligence; ml; syndrome differentiation; ai; bilstm-crf
7	logistic regression	8	2019	logistic regression; asthma; real-world evidence; syndrome; registry study
8	drug-drug interactions	8	2015	drug-drug interactions; national health insurance research; database; alert system; herb-drug interactions;
9	covid-19	8	2020	covid-19; clinical features; traditional Chinese medicine; electronic medical records; propensity score match

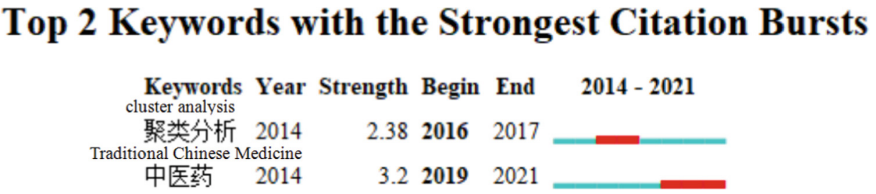

Fig. 8. Keyword Emergence Analysis of Chinese Literature

data analysis and mining, etc. From Figs. 2 and 3, we know that the density of author-cooperative network maps in the literature in this area is low. The cooperation density of Chinese and English authors was 0.0192 and 0.0374 respectively. In the future, it is necessary to strengthen cooperation and form a team of high-impact authors to promote high-quality development in the field of TCM electronic medical record research.

From the perspective of institutions of publication, the top 5 institutions of publication are all from TCM universities except China Academy of Chinese Medical Sciences. This shows that colleges and universities are the main institutions of TCM electronic medical record research. The China Academy of Chinese Medical Sciences ranks first in the number of publications in this field, but it is also a comprehensive TCM research institution integrating scientific research, medical treatment and teaching. In addition, the cities with the most publications are more economically developed. It can be seen that economic support and cultural exchange are also important driving forces for the development of TCM electronic medical record research.

Through keyword co-occurrence and cluster analysis, we found that scholars pay more attention to the standardization of TCM electronic medical record. As can be seen from Fig. 5, data mining on acupuncture points, combined drug use, big data and other application studies of TCM electronic medical record data have great influence and have been receiving attention. According to Fig. 6, Fig. 7, and Table 3, we know that the most prominent research hotspots in this field abroad are the optimization, design and application of algorithms such as deep learning, neural network, cluster analysis and association rules to assist the diagnosis and treatment of diseases based on the electronic medical record data of traditional Chinese medicine. It is noteworthy that data mining and artificial intelligence are common research focuses in this field at home and abroad. Correlation analysis, cluster analysis, complex network, logistic regression and other methods are widely used in China, while deep learning, neural network, Bayesian network, support vector machine and other methods are widely used abroad.

5 Conclusion

In recent years, more and more attention has been paid to the electronic medical record of Chinese medicine. The research focus has changed. The original research focus is the standardization of electronic medical records, information system construction and other basic research. The research hotspot after the change is the applied research on the construction of TCM intelligent assisted diagnosis and treatment system based on TCM electronic medical record and the data mining algorithm in line with TCM characteristics.

References

1. SHENG Hui, WANG Zhenguo. Design and Research on the Big Data Service Platform of Electronic Medical Recordof Traditional Chinese Medicine[J]. JOURNAL OF SHANDONG UNIVERSITY OF TCM, 2017, 41(5):402–405.
2. CHEN Yue, CHEN Chao-mei, LIU Ze-yuan, etc. The methodology function of Cite Space mapping knowledge domains[J]. Studies in Science of Science, 2015, 33(2):242–253.
3. WANG Xiaoxiao, GUO Qing. Recent 10-year Research Hotspots and Development Trends of Integrated Medical and Nursing Care in China: a CiteSpace-based Visual Analysis[J]. Chinese General Practice, 2021, 24(1):92–97.
4. CHEN Yao, ZENG Xue-yang, YANG Feng-wen, etc. Visual analysis of knowledge map of network Meta-analysis in traditional Chinese medicine based on CiteSpace[J]. China Journal of Chinese Materia Medica, 2020, 45(18):4500–4509.
5. LI Yu, YANG Tao. Analysis of the Current Situation and Development Trend of Data Mining of Traditional Chinese Medicine based on CiteSpace[J]. Guiding Journal of Traditional Chinese Medicine and Pharmacy, 2021, 27(04):153–157.
6. RONG Guang, XIE Qing-yu, MENG Qing-gang. Electronic Medical Record in TCM Field—An Analysis Based on Scientific Knowledge Visualization[J]. Chinese Journal of Information on Traditional Chinese Medicine, 2017, 24(1):99–104.
7. XIAO Yong, SHEN Shaowu, TIAN Shuanggui. Development Course and Prospect of Informatization Building of Traditional Chinese Medicine Hospitals in China[J]. Journal of Medical Informatics, 2021, 42(07):2–6+11.