

A Data-Driven Approach for Discovering the Recent Research Status of Diabetes in China

Xieling Chen¹, Heng Weng^{2(✉)}, and Tianyong Hao^{3(✉)}

¹ School of Economics, Jinan University, Guangzhou, China
shaylyn_chen@163.com

² The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, China
ww128@qq.com

³ School of Information Science and Technology, Guangdong University of Foreign Studies, Guangzhou, China
haoty@gdufs.edu.cn

Abstract. This paper aims at discovering the recent research status of diabetes in China through a data-driven bibliometrics and knowledge mapping analysis method on diabetes-related literature. With the basis of 24,561 publication documents from CNKI during 2007–2016, the quantitative analysis are conducted in three aspects: (1) descriptive statistical method for acquiring literature distribution characteristics; (2) hierarchical clustering, *k*-means clustering analysis, and multidimensional scaling analysis based on a keyword co-occurrence matrix for discovering research hotspots; and (3) network analysis for revealing cooperation relationships among authors and affiliations. The result shows some findings about the recent diabetes research in China. It also demonstrates the close cooperation of diabetes research among productive authors and affiliations through network generation and visualization.

Keywords: Diabetes · Distribution characteristics · Research hotspots · Cooperation network

1 Introduction

Diabetes and its complications are now the major health killers in most countries. According to International Diabetes Federation [1], 415 million or 8.8% of adults aged 20–79 had diabetes in 2015. The number is estimated to be 642 million in 2040. Not only does diabetes pose great threat to worldwide health, but it makes substantial economic impact on countries and national health systems. China is the top 1 country with the most people suffering from diabetes. According to Chinese Diabetes Society [2], there were nearly 110 million people with diabetes in China in 2014 with the overall prevalence doubling to 9.7% in nearly 10 years. The number of people at high risk was 150 million with the prevalence of pre-diabetes up to 15.5%, which was higher than the world average. Compared with the worldwide situation, threaten brought by diabetes is more severe in China.

The research of diabetes has been always of an extensive concern in academia. Doctors and researchers have published a wealth of diabetes literatures to record their clinical and research findings. These research literatures can to some extent reflect the development of diabetes prevention and treatment. However, it is difficult to accurately capture the pulse through traditional manual retrieval when facing with diabetes literatures of such a large quantity, wide distribution and fast growth. Therefore, it is of need to develop an automatic, efficient, and accurate method to continuously discover the recent research status of diabetes.

Bibliometrics uses relevant statistical and mathematical approaches to study information materials, which has become a well-established part of information research to the quantitative description of documents [3]. Traditional bibliometrics is popular in solving problems of extreme value and sorting, but it is weak in revealing the structure of literature, e.g., the cooperation among authors and affiliations, the structure and evolution of research hotspots. On the other hand, knowledge mapping analysis is a widely applied graphical and visual technique in revealing structure relationship and scientific knowledge development process [4]. Technologies involving in mathematics, information science, and computer science are combined to help analyze and understand the development process and forefront issues [5]. To a certain extent, knowledge mapping analysis and bibliometrics can be regarded as complements for each other.

This paper presents a data-driven method based on bibliometrics and knowledge mapping analysis, aiming at discovering the recent research status of diabetes in China. Using the techniques of descriptive statistics, clustering and multidimensional scaling, and network analysis, this paper conducts the analysis on literature distribution characteristics, research hotspots and cooperation relationships. We believe that the work can potentially assist clinical professionals and medical researchers in determining hot research topics and keeping abreast of the research status of diabetes in the development of research strategies.

2 Literature Review

There have been a considerable number of studies with the applications of bibliometrics and knowledge mapping analysis: evaluate the speed of publication of ophthalmology journals [6], map the literature related to a certain research field such as cancer research [7, 8] or health literacy research [9], and allow one to recognize new topics in the literature [10]. Bibliometrics and knowledge mapping analysis have played fundamental roles in examining the trends in medical research output. Ramos [11] conducted a bibliometric analysis of tuberculosis research indexed in PubMed during 1997 and 2006, finding the phenomenon that the research output in countries with more estimated cases of tuberculosis was less than that in industrialized countries. Boudry et al. [12] provided a review on scientific production related with the field of eye disease during 2010 and 2014 so as to identify the major topics as well as the predominant actors including journals, countries, and continents. Similar works have been conducted for other medical fields, such as obesity [13], leishmaniasis [14],

methaemoglobinaemia [15], giardiasis [16], dermatology [17], neurogenic bladder [18], rheumatology [19], and surgery [20].

Few studies concerning with diabetes using bibliometrics and knowledge mapping analysis methods can be found. Harande [21] used bibliometric approach to examine the increasing diabetes-related literature in Nigeria, indicating that the literature of diabetes in Nigeria was in harmony with the Bradford–Zipf distribution. Harande and Alhaji [22] examined the growth of published literature on the disease in three countries including Nigeria, Argentina and Thailand, showing that the literature of diabetes grew and spread very widely. Similar research has been done for diabetes literature in Middle East countries [23]. Geaney et al. [24] provided a detailed evaluation of type 2 diabetes mellitus research output during 1951–2012 with methods of large-scale data analysis, bibliometric indicators, and density-equalizing mapping. Zhang et al. [25] investigated the relationship between antipsychotics and type 2 diabetes research with bibliometrics method. The existing relative studies seldom focus on the literature of diabetes in China, let alone use the combination methods of bibliometrics and knowledge mapping analysis.

Therefore, the aim of this study is to provide a detailed evaluation of the diabetes research output from 2007 to 2016 in China using a data-driven method based on bibliometrics and knowledge mapping analysis to quantitatively analyze data from the CNKI database in terms of: (1) literature distribution characteristics with method of descriptive statistics; (2) research hotspots with method of clustering and multidimensional scaling analysis; and (3) cooperation relationships among authors and affiliations with network analysis methods.

3 The Data-Driven Approach

With the combination of bibliometrics and knowledge mapping analysis, we design a data-driven approach for discovering the recent research status of diabetes in mainland China from medical literature. The processing procedure is shown as Fig. 1. Diabetes-related publications as raw data are retrieved from CNKI and are used as dataset after preprocessing for bibliometrics and knowledge mapping analysis. The analysis applies descriptive statistical methods for acquiring literature distribution characteristics, multidimensional scaling analysis and data mining methods such as *k*-means clustering for discovering research hotspots, as well as network analysis for revealing cooperation relationships among authors and affiliations. Specifically, the procedure can be divided into the following stages:

Literature retrieval: The relevant publications were directly retrieved from CNKI in April 2017 with the following searching criteria: (1) “diabetes” as topic using an extended search keyword list; (2) “2007” to “2016” as publication year; (3) SCI, EI, CSSCI, and core journals as literature sources to keep high literature quality. The total retrieved raw data containing 35,059 documents was obtained for preprocessing.

Data preprocessing: to make sure of the high relevance with diabetes research, further data cleaning and processing was needed, consisting of: (1) excluding articles belonging to messaging literature like yearbooks, newspapers, and non-academic book publications; (2) excluding articles whose titles and keywords did not contain high

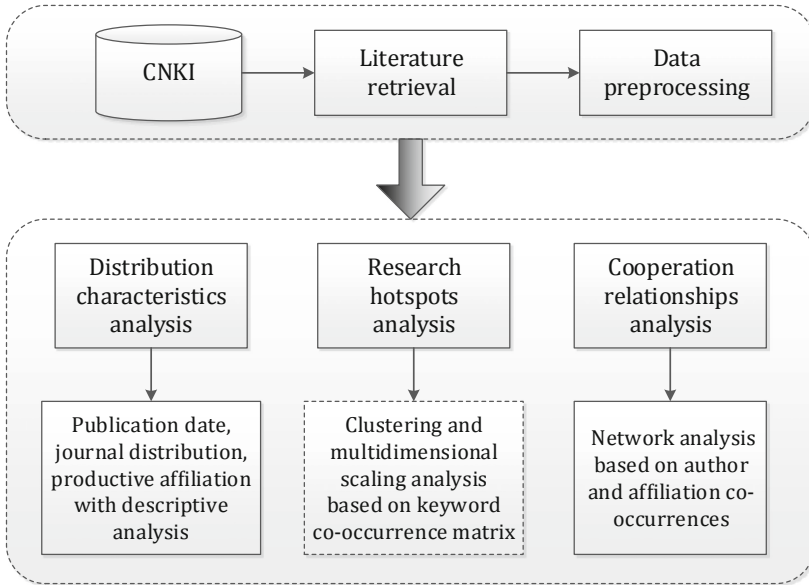


Fig. 1. The procedure for acquiring research status of diabetes in China

relevant words like “diabetes” and “hypoglycemic”; (3) keeping all the articles whose source journal names contained the keyword “diabetes”. After preprocessing, we finally acquired the dataset containing 24,561 articles.

Distribution characteristics analysis: with the basis of 24,561 documents, the literature distribution characteristics including publication date and journal distribution as well as productive affiliations were acquired using descriptive statistical methods. For example, as for productive affiliations, the affiliations were ranked according to their decreasing order of productivity. The rankings started with the affiliations with the highest number of diabetes-related publications. We used a publicly available software R for statistical analysis in the paper.

Research hotspots analysis: the recent research hotspots of diabetes in China were acquired based on a keyword co-occurrence matrix with the following steps. (1) Standardization of keywords: there existed nonstandard and inconsistent keyword representations thus it needed to be consistent, e.g., unifying “type 2 diabetes”, “diabetes, type 2”, “diabetes type 2”, “Type 2 diabetes (T2DM)” as “type 2 diabetes”; unifying “type 1 diabetes”, “diabetes, type 1”, “diabetes type 1”, “Type 1 diabetes (T1DM)” as “type 1 diabetes”; unifying “non-insulin-dependent diabetes mellitus”, “diabetes, non-insulin-dependent” as “non-insulin-dependent diabetes mellitus”; unifying “Type 1”, “TypeI” as “Type 1”; unifying “Type 2”, “TypeII” as “Type 2”; unifying “gestational diabetes”, “diabetes, pregnancy”, “diabetes, gestational” as “gestational diabetes”. (2) With the standardization, keywords with high frequency were ranked and sorted. (3) With the top frequent keywords, a co-occurrence matrix was constructed. (4) With the co-occurrence matrix, a keyword correlation matrix was acquired through the calculation using Ochiai correlation coefficient. (5) With the

keyword correlation matrix, a keyword dissimilar matrix was generated. (6) Hierarchical clustering and k -means clustering analysis were applied based on the keyword correlation matrix. (7) Multidimensional scaling analysis was implemented based on the keyword dissimilar matrix.

Cooperation network analysis: With the top productive authors and affiliations generated with descriptive statistical method, the cooperation network of authors and affiliations were acquired, respectively. It was required to convert the author and affiliation data into date form of one-to-one cooperation at first, and then the cooperation networks were generated using an easy-to-use package named as *networkD3* in R. It visualized the networks into force-directed graphs using the a function named as *simpleNetwork*.

4 Results and Discussion

As for literature characteristics analysis, the publication date and journal distribution as well as productive affiliations were taken into consideration. Table 1 shows the number and growth rate of diabetes publications by year during 2007–2016, revealing that the total number of publications increased rapidly by year before 2010, but dwindled after 2010. The growth rate of publications on diabetes from 2009 to 2010 reached up to 31.9%, witnessing the great mass upsurge on diabetes research in 2010.

Table 1. The number and growth rate of publications on diabetes by year during 2007–2016

Year	#publications	#cumulative publications	Growth rate (%)
2007	2,202	2,202	–
2008	2,066	4,268	–6.176
2009	2,326	6,594	12.585
2010	3,068	9,662	31.900
2011	2,768	12,430	–9.778
2012	2,337	14,767	–15.571
2013	2,433	17,200	4.108
2014	2,458	19,658	1.028
2015	2,589	22,247	5.330
2016	2,314	24,561	–10.622

The time regression curve of cumulative number of publications could be fitted as $y = -5.075 \times 10^6 + 2.529 \times 10^3 x$ with the goodness-of-fit R^2 equaling to 0.9991. It indicates that the regression curve fitted the development trend of the actual cumulative number of publications well, with which the future research output on diabetes in China could be inferred. For example, the predicted number of publications on diabetes research in 2017 in China is $-5.075 \times 10^6 + 2.529 \times 10^3 \times 2017 - (-5.075 \times 10^6 + 2.529 \times 10^3 \times 2016) = 2529$. Table 2 shows the top 10 most productive journals, from which these 10 journals together accounted for 37.52% of the total publications.

Table 2. Top 10 most productive journals

Journal	#publications	Proportion %	Cumulative proportion %
Chinese Journal of Diabetes	2,431	9.90	9.90
Chinese Journal of Gerontology	1,895	7.71	17.61
Shandong Medical Journal	964	3.92	21.54
Chinese General Practice	885	3.60	25.14
The Journal of Practical Medicine	563	2.29	27.43
Guangdong Medical Journal	520	2.12	29.55
Modern Preventive Medicine	518	2.11	31.66
China Journal of Modern Medicine	495	2.02	33.67
Chongqing Medicine	482	1.96	35.63
Lishizhen Medicine and Material Medical Research	463	1.88	37.52

Table 3 shows the top 10 most productive first author affiliations and total author affiliations, from which *Anhui Provincial Hospital* ranked within top two in both the top 10 most productive first author affiliations and total affiliations, indicating that it contributed much on the diabetes research in China.

Table 3. Top 10 most productive first author affiliations and total affiliations

Rank	First author affiliations	#publications	Rank	Total author affiliations	#publications
1	Anhui Provincial Hospital	298	1	Beijing University of Chinese Medicine	641
2	First Affiliated Hospital of Chongqing Medical University	212	2	Anhui Provincial Hospital	618
3	General Hospital of PLA	202	3	Guangzhou University of Chinese Medicine	453
4	Shanghai Jiao Tong University Affiliated Sixth People's Hospital	200	4	Shanghai Jiao Tong University Affiliated Sixth People's Hospital	421
5	Nanjing Medical University	187	5	Nanjing Medical University	392
6	Beijing University of Chinese Medicine	177	6	General Hospital of PLA	380
7	First Affiliated Hospital, Nanjing Medical University	160	7	Peking University First Hospital	332
8	Peking University First Hospital	150	8	First Affiliated Hospital of Chongqing Medical University	331
9	Guangzhou University of Chinese Medicine	149	9	Chengdu University of TCM	321
10	Chengdu University of TCM	146	10	First Affiliated Hospital, Nanjing Medical University	276

For the research hotspots analysis, we used author defined keywords. Generally, they can represent the main points of the publications. Table 4 shows 34 keywords with frequency greater than or equaling to 200. The top 3 in order were: “Diabetes” with frequency of 7,510 and frequency of 7.078%, “Type 2 Diabetes” with frequency of 6,978 and frequency of 6.577%, and “Diabetic Nephropathy” with frequency of 2,873 and frequency of 2.708%. Using the keywords, a co-occurrence matrix with 34 rows and 34 columns was generated, where the top 10 keywords are shown as Table 5. The co-occurrence matrix is a symmetric matrix. The data on the main diagonal indicates the frequency of the keywords and the data on the non-main diagonal represents the co-occurrence frequency between two different keywords. For instance, the co-occurrence frequency of “Diabetes” and “Blood Glucose” is 280, indicating that these two keywords appeared together in 280 documents.

Table 4. The ranked keywords with frequency greater than or equaling to 200

Rank	Keyword	Frequency	Frequency %	Rank	Keyword	Frequency	Frequency %
1	Diabetes	7,510	7.078	18	Blood Fat	352	0.332
2	Type 2 Diabetes	6,978	6.577	19	Coronary Heart Disease	321	0.303
3	Diabetic Nephropathy	2,873	2.708	20	Diabetes Rat	315	0.297
4	Insulin Resistance	1,372	1.293	21	Obesity	313	0.295
5	Gestational Diabetes	949	0.894	22	Metformin	283	0.267
6	Blood Glucose	847	0.798	23	C-reactiveprotein	276	0.26
7	Diabetic Retinopathy	752	0.709	24	Fasting Plasma Glucose	266	0.251
8	Diabetic Peripheral Neuropathy	717	0.676	25	Genetic Polymorphism	257	0.242
9	Insulin	625	0.589	26	Acupuncture	252	0.238
10	Oxidative Stress	624	0.588	27	Atherosclerosis	245	0.231
11	Diabetic Foot	595	0.561	28	Streptozotocin	241	0.227
12	Risk Factors	581	0.548	29	Meta Analysis	226	0.213
13	Rat	574	0.541	30	Hypoglycemic	222	0.209
14	Hypertension	522	0.492	31	Metabolic Syndrome	216	0.204
15	Glycosylated Hemoglobin	467	0.44	32	Vascular Endothelial Growth Factor	205	0.193
16	Type 1 Diabetes	407	0.384	33	Diabetic Cardiomyopathy	202	0.19
17	Adiponectin	364	0.343	34	Transforming Growth Factor-β1	201	0.189

After that, we generated a keyword correlation matrix with the basis of the co-occurrence matrix using Ochiai correlation coefficient. The used calculation formula was $O_{ij} = A_{ij} / \sqrt{A_i A_j}$. In the formula, the value range of O_{ij} was [0, 1], representing the probability of the co-occurrence of keyword W_i and W_j . A_{ij} represented the co-occurrence frequency of keyword W_i and W_j . A_i represented the frequency of the keyword W_i and A_j represented the frequency of the keyword W_j . In the correlation matrix, the value represented the distance between two keywords. The larger the

Table 5. The top 10 keywords in the co-occurrence matrix

	Blood Glucose	Diabetes	Diabetic Nephropathy	Diabetic Peripheral Neuropathy	Diabetic Retinopathy	Gestational Diabetes	Insulin	Insulin Resistance	Oxidative Stress	Type2 Diabetes
Blood Glucose	847	280	36	7	4	38	101	36	14	273
Diabetes	280	7,510	109	96	53	6	223	168	143	10
Diabetic Nephropathy	36	109	2,873	3	23	2	10	63	146	235
Diabetic Peripheral Neuropathy	7	96	3	717	2	0	1	2	39	103
Diabetic Retinopathy	4	53	23	2	752	1	1	8	27	92
Gestational Diabetes	38	6	2	0	1	949	43	131	9	4
Insulin	101	223	10	1	1	43	625	33	7	204
Insulin Resistance	36	168	63	2	8	131	33	1,372	40	635
Oxidative Stress	14	143	146	39	27	9	7	40	624	134
Type2 Diabetes	273	10	235	103	92	4	204	635	134	6,978

correlation value was between two keywords, the smaller the distance was between them. Then the keyword dissimilar matrix was acquired through subtracting each value in the correlation matrix from 1.

Therefore, 34 keywords were divided into 7 research hotspot categories through hierarchical clustering, *k*-means clustering, and multidimensional scaling analysis, respectively. We set the cluster number *k* as 7 in the experiment from our empirical experience. Figure 2 shows the result of hierarchical clustering using complete-linkage with distance function $D_{HK} = \max(d_{uv}), u \in H$ and $v \in K$. d_{uv} represented the distance between word *u* and *v*, where *u* belonged to cluster *H* and *v* belonged to cluster *K*.

Among the generated research hotspot categories using the three methods, three categories are the same as follows: (1) Glycosylated Hemoglobin, and Type 2 Diabetes; (2) Blood Fat, and Blood Glucose; and (3) Diabetic Retinopathy, Vascular Endothelial Growth Factor, Diabetic Nephropathy, and Transforming Growth Factor-β1. Table 6 is a summary of the results of hierarchical clustering, *k*-means clustering, and multidimensional scaling analysis.

For cooperation network analysis, the number of publications with a specific first author or affiliation was much less than that with cooperation with other authors or affiliations. For example, the number of publications for *Anhui Provincial Hospital* as first author affiliation was 298. However, the number was up to 618 when cooperating with other affiliations, indicating the cooperation among authors or affiliations on diabetes research was much. Therefore, considering all authors in the same publications, we analyzed the cooperation relationships among authors with the number of

Table 6. The summary of the results of hierarchical clustering, *k*-means clustering, multidimensional scaling analysis

Category	Hierarchical clustering	<i>k</i> -means clustering	Multidimensional scaling analysis
1	Glycosylated Hemoglobin; Type 2 Diabetes	Glycosylated Hemoglobin; Type 2 Diabetes	Glycosylated Hemoglobin; Type 2 Diabetes
2	Blood Fat; Blood Glucose	Blood Fat; Blood Glucose	Blood Fat; Blood Glucose
3	Diabetic Retinopathy; Vascular Endothelial Growth Factor; Diabetic Nephropathy; Transforming Growth Factor-β1	Diabetic Retinopathy; Vascular Endothelial Growth Factor; Diabetic Nephropathy; Transforming Growth Factor-β1	Diabetic Retinopathy; Vascular Endothelial Growth Factor; Diabetic Nephropathy; Transforming Growth Factor-β1
4	Metabolic Syndrome; Obesity; Risk Factors; Coronary Heart Disease; Diabetes; Hypertension; Atherosclerosis; C-reactiveprotein	Metabolic Syndrome; Obesity; Risk Factors; Coronary Heart Disease; Diabetes; Hypertension; Atherosclerosis	Metabolic Syndrome; Obesity; Risk Factors; Coronary Heart Disease; Diabetes; Hypertension; Atherosclerosis; Gestational Diabetes; Type 1 Diabetes; Hypoglycemic; Adiponectin; Insulin Resistance; Insulin
5	Genetic Polymorphism; Meta Analysis; Acupuncture; Diabetic Foot; Fasting Plasma Glucose; Diabetes Rat; Oxidative Stress; Hypoglycemic; Streptozotocin; Type 1 Diabetes;	Genetic Polymorphism; Meta Analysis; Acupuncture; Diabetic Foot; Insulin; Diabetic Peripheral Neuropathy; Metformin; Type 1 Diabetes;	Genetic Polymorphism; Meta Analysis; Acupuncture; Diabetic Foot; Diabetes Rat; Oxidative Stress; Streptozotocin; Diabetic Cardiomyopathy; Fasting Plasma Glucose; C-reactiveprotein; Metformin; Rat
6	Gestational Diabetes; Adiponectin; Insulin Resistance; Insulin; Metformin; Diabetic Cardiomyopathy; Rat	Gestational Diabetes; Adiponectin; Insulin Resistance; C-reactiveprotein	Hypertension
7	Diabetic Peripheral Neuropathy	Hypoglycemic; Fasting Plasma Glucose; Oxidative Stress; Rat; Diabetes Rat; Streptozotocin; Diabetic Cardiomyopathy	Diabetic Peripheral Neuropathy

publications greater than or equaling to 100, 150, and 200, respectively, as well as the affiliations with the number of publications greater than or equaling to 50.

Figure 3 shows the cooperation network of authors with the number of publications ≥ 100 (access via the link¹). Figure 4 shows the cooperation network of affiliations with the number of publications ≥ 50 (access via the link²). In the network, the black

¹ http://www.zhukun.org/haoty/resources.asp?id=HIS2017_author_100.

² http://www.zhukun.org/haoty/resources.asp?id=HIS2017_affiliation_50.

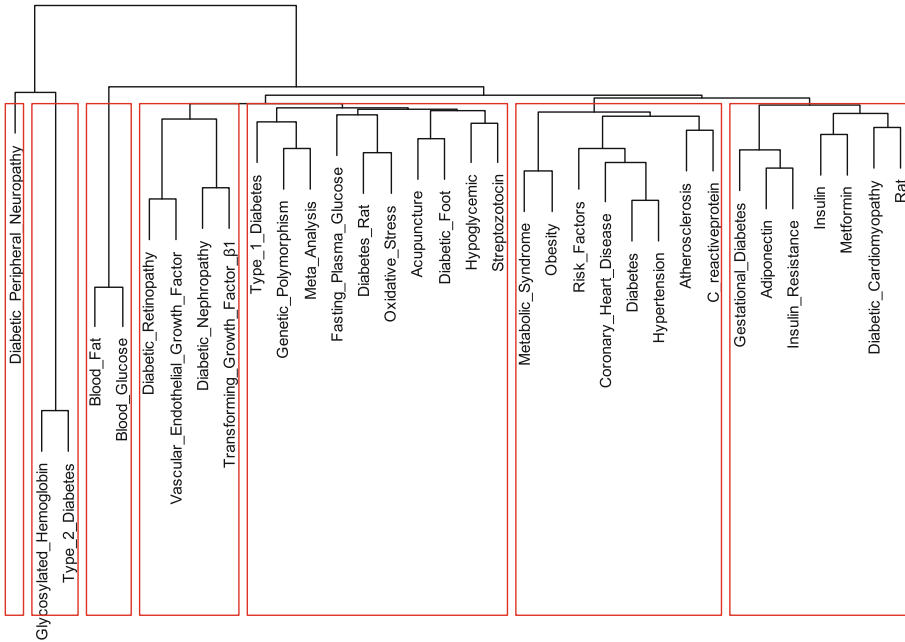


Fig. 2. The visualized result of hierarchical clustering

nodes represent authors or affiliations, and the lines represent the cooperation relationship. The more connected lines surrounding a specific node, the closer cooperation relationship with other authors or affiliations is for the author or affiliation. Moreover, one can dynamically drag and drop to view the cooperation relationship for a specific author or affiliation.

5 Summary

This paper presented a data-driven method based on bibliometrics and knowledge mapping analysis, aiming at discovering the recent research status of diabetes in China during 2007–2016. With the analysis techniques of descriptive statistics, clustering and multidimensional scaling, and network analysis, we acquired literature distribution characteristics, research hotspots, and cooperation relationships among authors and affiliations. The results and findings were presented. Our work can provide clinical researchers and funding agencies with state-of-the-art research status, potentially assisting scientific research topic determination and participating in the development of clinical research strategies.



Fig. 3. The cooperation network of authors with the number of publications greater than or equaling to 100

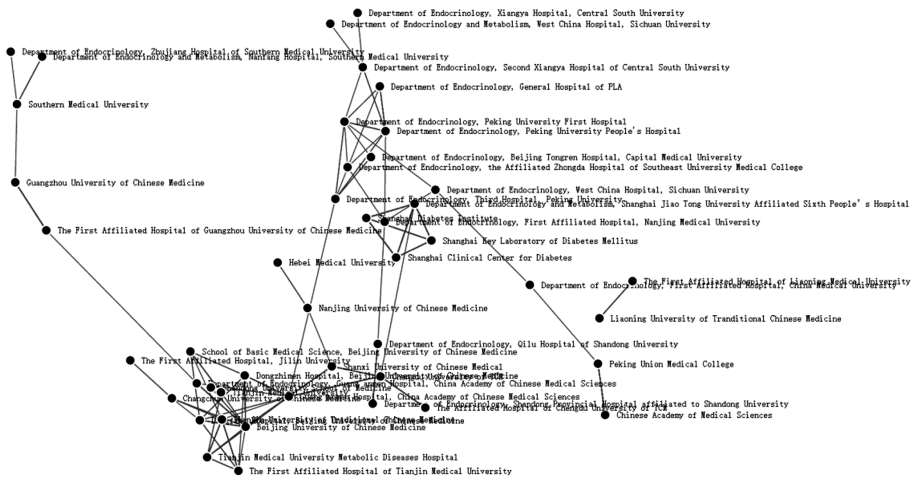


Fig. 4. The cooperation network of affiliations with the number of publications greater than or equaling to 50

Acknowledgements. The work was supported by National Natural Science Foundation of China (No. 61772146 & 61403088), Frontier and Key Technology Innovation Special Grant of Guangdong Province (No. 2014B010118005), Public Interest Research and Capability Building Grant of Guangdong Province (No. 2014A020221039), and Innovative School Project in Higher Education of Guangdong Province (No. YQ2015062).

References

1. International Diabetes Federation (IDF): IDF Diabetes Atlas Seventh Edition (2016)
2. Chinese Diabetes Society: Guidelines for the Prevention and Treatment of Type 2 Diabetes in China (2013). *Chin. J. Diabetes Mellitus* 6(7) (2014)
3. Zhang, H.Q.: A bibliometric study on medicine chinese traditional in medline database. *Scientometrics* 31(3), 241–250 (1994)
4. Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., Pellegrino, D.: Towards an explanatory and computational theory of scientific discovery. *J. Infor.* 3(3), 191–209 (2009)
5. Guo, Y., Wang, X.M., Wei, H.E.: Chinese and international dynamic research into low carbon technology based on bibliometrics and knowledge mapping. *Inf. Sci.* 33(4), 139–148 (2015)
6. Chen, H., Chen, C.H., Jhanji, V.: Publication times, impact factors, and advance online publication in ophthalmology journals. *Ophthalmology* 120(8), 1697–1701 (2013)
7. Holliday, E.B., Ahmed, A.A., Yoo, S.K., Jagsi, R., Hoffman, K.E.: Does cancer literature reflect multidisciplinary practice? A systematic review of oncology studies in the medical literature over a 20-year period. *Int. J. Rad. Oncol. Biol. Phys.* 92(4), 721–731 (2015)
8. Salisu, S.A., Ojoye, O.T.: Bibliometric analysis of cancer publications in Nigeria during 2008–2012. *Int. J. Libr. Inf. Sci.* 7(3), 69–76 (2015)
9. Kondilis, B.K., Kiriaze, I.J., Athanasoulia, A.P., Falagas, M.E.: Mapping health literacy research in the European Union: a bibliometric analysis. *PLoS ONE* 3(6), e2519 (2008)
10. Zacca-González, G., Chinchilla-Rodríguez, Z., Vargas-Quesada, B., de Moya-Anegón, F.: Bibliometric analysis of regional Latin America's scientific output in public health through SCImago journal & country rank. *BMC Publ. Health* 14(1), 632 (2014)
11. Ramos, J.M., Padilla, S., Masia, M., Gutierrez, F.: A bibliometric analysis of tuberculosis research indexed in PubMed, 1997–2006. *Int. J. Tuberculosis Lung Disease* 12, 121461–121468 (2008)
12. Boudry, C., Denion, E., Mortemousque, B., Mouriaux, F.: Trends and topics in eye disease research in PubMed from 2010 to 2014. *PeerJ* 4, e1557 (2016)
13. Khan, A., Choudhury, N., Uddin, S., Hossain, L., Baur, L.A.: Longitudinal trends in global obesity research and collaboration: a review using bibliometric metadata. *Obes. Rev.* 17(4), 377–385 (2016)
14. Perilla-González, Y., Gómez-Suta, D., Delgado-Osorio, N., Hurtado-Hurtado, N., Baquero-Rodriguez, J.D., Lopez-Isaza, A.F., Lagos-Grisales, G.J., Villegas, N., Rodriguez-Morales, A.: Study of the scientific production on leishmaniasis in Latin America. *Recent Patents Anti Infective Drug Disc.* 9(3), 216–222 (2014)
15. Sa'ed, H.Z., Al-Jabi, S.W., Sweileh, W.M., Al-Khalil, S., Alqub, M., Awang, R.: Global methaemoglobinaemia research output (1940–2013): a bibliometric analysis. *Springerplus* 4(1), 626 (2015)
16. Escobedo, A.A., Arencibia, R., Vega, R.L., Rodríguez-Morales, A.J., Almirall, P., Alfonso, M.: A bibliometric study of international scientific productivity in giardiasis covering the period 1971–2010. *J. Infection Developing Countries* 9(1), 076–086 (2015)

17. Man, H., Xin, S., Bi, W., Lv, C., Mauro, T.M., Elias, P.M., Man, M.Q.: Comparison of publication trends in dermatology among Japan, South Korea and Mainland China. *BMC Dermatol.* **14**(1), 1 (2014)
18. Gao, Y., Qu, B., Shen, Y., Su, X.J., Dong, X.Y., Chen, X.M., Pi, H.Y.: Bibliometric profile of neurogenic bladder in the literature: a 20-year bibliometric analysis. *Neural Regener. Res.* **10**(5), 797 (2015)
19. Cheng, T., Zhang, G.: Worldwide research productivity in the field of rheumatology from 1996 to 2010: a bibliometric analysis. *Rheumatology* **52**(9), 1630–1634 (2013). ket008
20. Sharma, B., Boet, S., Grantcharov, T., Shin, E., Barrowman, N.J., Bould, M.D.: The h-index outperforms other bibliometrics in the assessment of research performance in general surgery: a province-wide study. *Surgery* **153**(4), 493–501 (2013)
21. Harande, Y.I.: Exploring the literature of diabetes in Nigeria: a bibliometrics study. *Afr. J. Diabetes Med.* **19**(2), 1–4 (2011)
22. Harande, Y.I., Alhaji, I.U.: Basic literature of diabetes: a bibliometrics analysis of three countries in different world regions. *J. Libr. Inf. Sci.* **2**(1), 49–56 (2014)
23. Nasli-Esfahani, E., Farzadfar, F., Kouhnavard, M., Ghodssi-Ghassemabadi, R., Khajavi, A., Peimani, M., Sanjari, M.: Iran diabetes research roadmap (IDRR) study: a preliminary study on diabetes research in the world and Iran. *J. Diabetes Metabolic Disorders* **16**(1), 9 (2017)
24. Geaney, F., Scutaru, C., Kelly, C., Glynn, R.W., Perry, I.J.: Type 2 diabetes research yield, 1951–2012: bibliometrics analysis and density-equalizing mapping. *PLoS ONE* **10**(7), e0133009 (2015)
25. Zhang, Y., Shen, X., Chen, D.: Bibliometrics analysis of the relationship research of antipsychotics and type 2 diabetes. *Chin. J. Drug Dependence* **1**, 19 (2011)