*Computers, Materials & Continua*
DOI:10.32604/cmc.2022.017295
***Article***

Tech Science Press

# Traditional Chinese Medicine Automated Diagnosis Based on Knowledge Graph Reasoning

**Dezheng Zhang[1,2], Qi Jia[1,2], Shibing Yang[1,2], Xinliang Han[2], Cong Xu[3], Xin Liu[1,4] and Yonghong Xie[1,2,*]**

[1]School of Computer & Communication Engineering, University of Science & Technology Beijing, Beijing, 100083, China
[2]Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing, 100083, China
[3]Inspur Electronic Information Industry Co., Ltd. & State Key Laboratory of High-End Server & Storage Technology, Jinan, 250101, China
[4]Surgical Simulation Research Lab, Department of Surgery, University of Alberta, Edmonton, T6G 2E1, Alberta, Canada
[*]Corresponding Author: Yonghong Xie. Email: xieyh@ustb.edu.cn
Received: 26 January 2021; Accepted: 01 March 2021

**Abstract:** Syndrome differentiation is the core diagnosis method of Traditional Chinese Medicine (TCM). We propose a method that simulates syndrome differentiation through deductive reasoning on a knowledge graph to achieve automated diagnosis in TCM. We analyze the reasoning path patterns from symptom to syndromes on the knowledge graph. There are two kinds of path patterns in the knowledge graph: one-hop and two-hop. The one-hop path pattern maps the symptom to syndromes immediately. The two-hop path pattern maps the symptom to syndromes through the nature of disease, etiology, and pathomechanism to support the diagnostic reasoning. Considering the different support strengths for the knowledge paths in reasoning, we design a dynamic weight mechanism. We utilize Naïve Bayes and TF-IDF to implement the reasoning method and the weighted score calculation. The proposed method reasons the syndrome results by calculating the possibility according to the weighted score of the path in the knowledge graph based on the reasoning path patterns. We evaluate the method with clinical records and clinical practice in hospitals. The preliminary results suggest that the method achieves high performance and can help TCM doctors make better diagnosis decisions in practice. Meanwhile, the method is robust and explainable under the guide of the knowledge graph. It could help TCM physicians, especially primary physicians in rural areas, and provide clinical decision support in clinical practice.

**Keywords:** Traditional Chinese medicine; automated diagnosis; knowledge graph; Naïve Bayes; syndrome differentiation

## 1 Introduction

As a complementary field of medicine outside the modern medicine system, traditional Chinese medicine (TCM) has played a significant role in the healthcare of China for thousands of years [1–4]. According to the China Public Health Statistical Yearbook, over 1 billon TCM treatments are carried out in China each year [5]. Syndrome differentiation is a core diagnosis method of TCM. It analyzes the specific pattern of symptoms, etiology, nature, and location of a disease and guides treatment strategies [6]. In TCM, syndrome is a concept that abstracts a set of symptoms and determines the phase of a disease [7]. Master TCM syndrome differentiation is an intricate and time-consuming process. Because syndrome differentiation is very complicated to conduct, it can be difficult to maintain stable efficacy when treating a given disease. Moreover, the number of TCM doctors cannot support the huge demand for TCM treatments.

In recent years, automated diagnosis has received much attention. Automated diagnosis systems utilizing artificial intelligence aim to diagnose and make decisions based on a patient's condition. Most reported research has applied artificial intelligence in modern medicine [8–10]. Automated diagnosis in TCM is more challenging. Some researchers have begun to study the application of information technology in TCM diagnosis [11,12]. Wang et al. [13] used raw free-text as original input and employed the naïve Bayes and the support vector machine classifier for automated diagnosis in TCM. Xu et al. [14] designed an artificial neural network as a classifier for syndrome differentiation and achieved good performance in diagnosing chronic obstructive pulmonary disease. Liu et al. [15] focused on lung cancer syndrome differentiation. They treated syndrome differentiation as a multilabel text classification task and utilized deep learning to model the clinical record text feature for classification. They also used a fusion model approach to obtain better performance than a single model. Meanwhile, Zhang et al. [16] developed a TCM assistive diagnostic system based on artificial intelligence. A long-short term memory network (LSTM) with a conditional random field (CRF) framework extracted features from raw medical record text. Then a convolutional neural network (CNN) was used to predict the disease type. Despite these TCM automated diagnosis systems having positive preliminary results, there are some limitations. The existing methods require a large volume of annotated clinical records for training. Furthermore, these methods lack interpretability for the diagnosis process. In practice, clinicians need an automated diagnosis method that does not rely on a large number of annotated data and is explainable.

Knowledge graphs may address these limitations. Knowledge graphs describe concepts, entities, events, and their relationships in the real world. The knowledge graph is the foundational knowledge resource used to implement artificial intelligence systems [17,18]. In TCM, the knowledge graph could organize fragmented theoretical knowledge. In this way, we could reinforce the connectivity of TCM knowledge and support the automated diagnosis method. Xie et al. [19] proposed a personalized diagnostic pattern mining method based on the TCM knowledge graph with a specific doctor's clinical records. Meanwhile, Yu et al. [20] and Zheng et al. [21] described the construction of a TCM knowledge graph using databases and documents. Zhang et al. [22] introduced a TCM knowledge graph based on ontology. Lastly, Xie et al. [23] proposed a TCM auxiliary diagnosis method combining a knowledge graph and reinforcement learning.

In this study, we propose an artificial intelligence TCM automated diagnosis method. This method simulates syndrome differentiation through deductive reasoning on a knowledge graph and infers syndromes from the patient's symptoms. We analyze the reasoning path patterns from symptom to syndromes on the knowledge graph. According to these patterns, we illuminate the inference process from a set of symptoms to syndromes with naïve Bayes. The proposed method reasons the syndrome results by calculating the possibility according to the weighted score of the path in the knowledge graph. We evaluate the performance of our method with real-world record sets and prove its effectiveness and practicality.

## 2 Method

### *2.1 Task Definition*

For a given symptom set $Sym = \{sym_1, sym_2, \ldots, sym_n\}$, where $sym_i$ is a symptom, and a given syndrome set $Sd = \{sd_1, sd_2, \ldots, sd_m\}$, which is pre-defined by a specific disease and wherein $sd_j$ is a syndrome, we infer the target syndrome utilizing the TCM knowledge graph of Zhang et al. [22]. For each syndrome, $P(sd_j \mid Sym\}$ represents the probability of syndrome $sd_j$ being in symptom set $Sym$. The inference process simulates syndrome differentiation and treats knowledge paths (or reason paths) in the knowledge graph as evidence for the inference. These paths need to be consistent with cognitive diseases in TCM and indicate the diagnosis decision-making process of physicians. We limit the length of the pattern to 2 because we believe these patterns could provide evidence for diagnosis. Therefore, we define the meta-path as the reasoning path pattern as in Tab. 1.

**Table 1:** Meta-path of knowledge graph for syndrome differentiation

| Hop length | Meta-path pattern |
|---|---|
| One-hop pattern | Symptom → syndrome |
| Two-hop pattern | Symptom → nature of disease → syndrome |
| | Symptom → etiology → syndrome |
| | Symptom → sathomechanism → syndrome |

### *2.2 Naïve Bayes Automated Diagnosis on Knowledge Graph*

In this section, we describe the automated diagnosis method. According to the definition of the task, the core question is the calculation of the probability $P(sd_j \mid Sym)$. Based on the Bayes formula, we can get this relation:

$$P(sd_j \mid Sym) \propto P(Sym \mid sd_j), \tag{1}$$

where $P(Sym \mid sd_j)$ represents the probability that the symptom set $Sym$ occurs in the condition of the syndrome $sd_j$, and $P(sd_j)$ is an priori probability and represents the possibility of syndrome $sd_j$ being the specific disease. $P(sd_j)$ can be defined by a TCM expert or calculated according to past medical records.

We consider that each symptom $sym_i$ in symptom set $Sym$ is independent. Therefore, we can obtain

$$P\left(Sym \mid sd_j\right) = P\left(sym_1, sym_2, \ldots, sym_n \mid sd_j\right) = \prod_{i=1}^{n} P(sym_i \mid sd_j). \tag{2}$$

Combining this with Eq. (1), we can obtain

$$P(sd_j \mid Sym) \propto P(sd_j) \prod_{i=1}^{n} P(sym_i \mid sd_j). \tag{3}$$

We define the inference score as in Eq. (4). In practice, we use log to avoid the result of the series multiplication being too small.

$$score = \log P(sd_j) + \sum_{i=1}^{n} \log P(sym_i \mid sd_j) \tag{4}$$

Next, we need to calculate $P(sym_i \mid sd_j)$. The knowledge path on the knowledge graph is the main reasoning principle. There are two kinds of path patterns: one-hop and two-hop. We define $f_1(\cdot)$ as the score function of the one-hop path and $f_2(\cdot)$ as the score function of the two-hop path.

For the one-hop path, we search all knowledge paths from each symptom $sym_i$ to every syndrome $sd_j$ and calculate the one-hop score as follows:

$$f_1\left(sym_i, sd_j\right) = \frac{c_{i,j}^{1}}{\sum_{j=1}^{m} c_{i,j}^{1}}, \tag{5}$$

where $c_{i,j}^{1}$ represents the number of one-hop knowledge paths from symptom $sym_i$ to syndrome $sd_j$.

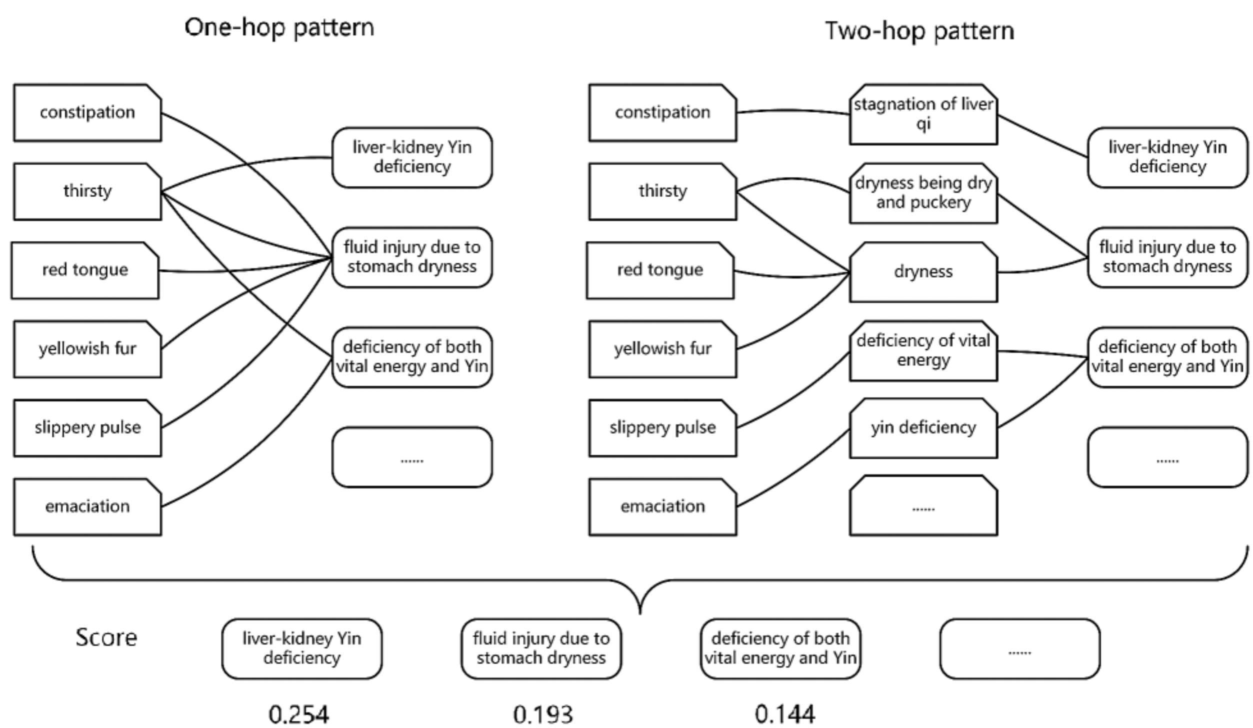


**Figure 1:** An example of automated diagnosis on the TCM knowledge graph

Two-hop paths represent the support from different perspectives in TCM, including nature of disease, etiology, and pathomechanism. However, the support strengths of the intermediate knowledge nodes for different syndromes are unequal. We use TF-IDF to regularize the path weight. As with the one-hop path, we first need to search all of the knowledge paths. Then, we calculate the path weight based on the frequency of each intermediate knowledge node for each different syndrome. The two-hop score is as follows:

$$f_2\left(sym_i, sd_j\right) = \frac{\sum_{k=1}^{c_{i,j}^2} \log\left(\frac{c_k^2}{m}\right)}{\sum_{j=1}^{m} c_{i,j}^2}, \tag{6}$$

where $c_{i,j}^2$ is the number of two-hop knowledge paths from symptom $sym_i$ to syndrome $sd_j$, and $c_k^2$ is the number of all the intermediate knowledge node paths. In this way, some intermediate knowledge nodes will be emphasized if they are particularly strongly related to a specific syndrome.

$P(Sym \mid sd_j)$ is determined by adding $f_1(\cdot)$ and $f_2(\cdot)$. Here, we use two hyper-parameters to balance the two different scores:

$$P\left(sym_i \mid sd_j\right) = \beta_1 f_1\left(sym_i, sd_j\right) + \beta_2 f_2(sym_i, sd_j). \tag{7}$$

We set $\beta_1 = 0.6$ and $\beta_2 = 0.4$ since we think a short path in the knowledge graph would provide better support than a long path. An example of the inference is shown in Fig. 1.

## 3 Experiments

### *3.1 Data Description*

We used two datasets to test our method. The first dataset was a clinical record set with data collected from the book, *Chinese Medical Records of All Famous Doctors*. We selected 519 clinical records related to nine different diseases, including coronary heart disease, diabetes, and some gynecological diseases. The medical records were manually processed by TCM experts. The syndromes of each disease were also defined by TCM experts. Tab. 2 lists the syndromes of each disease. The Chinese–English translations of the syndromes are presented in **Appendix A**.

We also used a real-world dataset. We deployed our method in nine hospitals to test its efficacy in practice. The hospitals included Guanganmen Traditional Chinese Medicine Hospital and Dongzhimeng Traditional Chinese Medicine Hospital in Beijing, China, among others. Doctors of these hospitals used our method to diagnose coronary heart disease and diabetes. Finally, doctors evaluated the result of the automated diagnosis based on their professional expertise. The distributions of gender, age, and syndromes are shown in Figs. 2–4, respectively.

### *3.2 Experiment Results*

We used the metrics Hit@N and MeanRank to evaluate the performance of the proposed method on the clinical record dataset. First, we ranked the list of predicted syndromes in descending order based on the possibility of correct reasoning. The Hit@N measures the probability of how often the correct syndrome is in the top N places of the list. Here, we set N to 1, 3, and 5. The MeanRank measures the average sorted position of the correct syndrome. For candidate set ranking, the aim is to rank the correct syndrome at the top position.

The performance metrics obtained for experiments on the clinical record dataset are shown in Tab. 3. The count column represents the number of clinical records pertaining to each disease. Our method gives the performance with Hit@1, Hit@3, and Hit@5 of 0.708, 0.958, and 0.980 and MeanRank of 1.438. Although coronary heart disease and diabetes have worse Hit@1 scores than other diseases, the Hit@5 is greater than 0.90 for both. As indicated by the results, the proposed method achieves high diagnostic accuracies on the eight diseases.

**Table 2:** The syndromes of each disease

| Disease | Syndromes |
|---|---|
| Coronary heart disease | Cold stagnation in heart vessel, cold stagnation and blood stasis, blood stasis due to Qi deficiency, deficiency of both vital energy and Yin, Qi stagnation in heart and chest, Qi depression to blood stasis, toxic-heat and blood stasis, turbid phlegm blocking, turbid phlegm and blood stasis, deficiency of heart Qi, heart-kidney Yang deficiency, heart-kidney Yin deficiency, heart blood stasis, Yang prostration |
| Diabetes | Stagnation of liver Qi and stomach heat, combination of phlegm and heat, consumption of body fluid due to lung-heat, fluid injury due to stomach dryness, fire excess from Yin deficiency, deficiency of both vital energy and Yin, liver-kidney Yin deficiency, deficiency of both Yin and Yang, floating of Yang due to Yin deficiency, phlegmatic hygrosis, stagnation of blood |
| Acyesis | Kidney-Yang deficiency, stagnation of liver-Qi, kidney-Yin deficiency, retention of phlegm-dampness in the interior, stagnation in uterus, kidney Qi deficiency, kidney deficiency and liver-Qi stagnation, Qi depression to blood stasis, cold-dampness to freeze and sluggish, intermingled phlegm and blood stasis, deficiency of the spleen causing stagnation of phlegm, damp-heat gluing, blood stasis and kidney deficiency |
| Uterine bleeding | Stagnation in uterus, kidney-Yin deficiency, asthenia pyrosyndrome, kidney-Yang deficiency, splenasthenic syndrome, excess-heat syndrome, kidney Qi deficiency |
| Menoxenia | Stagnation in uterus, deficiency of spleen-QI, kidney Qi deficiency, Yin deficiency and blood heat, Yang deficiency and blood heat, deficiency of liver-blood, stagnation of liver-Qi, phlegm stagnation, damp-heat gluing, pathogenic cold coagulating with blood, kidney Yang deficiency, stagnation of liver Qi and blood stasis, kidney deficiency and liver-Qi stagnation, insufficiency of both the spleen and the kidney, insufficiency of both the spleen and the kidney, insufficiency of vital energy and blood, blood stasis and kidney deficiency, deficiency of liver and kidney, Yin-deficiency and blood-dryness, deficiency of the kidney causing stagnation of phlegm, deficiency of the kidney and blood |
| Menorrhalgia | Qi depression to blood stasis, cold stagnation and blood stasis, dampness and heat stasis, deficiency of the liver and kidney, Qi-blood deficiency, blood stasis and kidney deficiency |
| Menostasis | Cold stagnation and blood stasis, phlegm stagnation, Qi depression to blood stasis, Qi-blood deficiency, kidney Qi deficiency, kidney-Yin deficiency, Yin asthenia generating intrinsic heat, blood and essence asthenia, Yin-deficiency and blood-dryness, liver-kidney Yin deficiency, splenasthenic syndrome, kidney-Yang deficiency |
| Menopausal syndrome | Failure of the heart and kidney integrating, kidney Yin and Yang deficiency kidney-Yang deficiency, kidney-Yin deficiency, fire excess from Yin deficiency, kidney deficiency and liver-Qi stagnation |

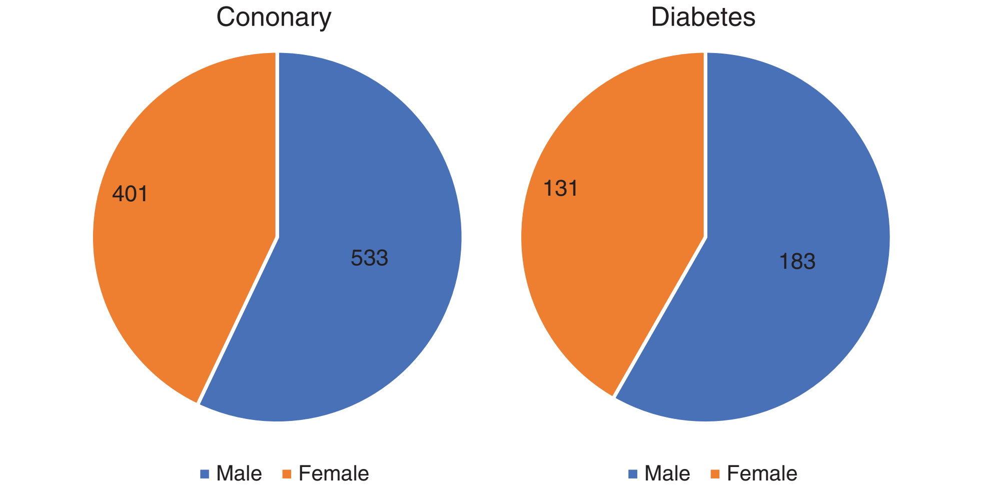


**Figure 2:** The distribution of gender for coronary heart disease and diabetes

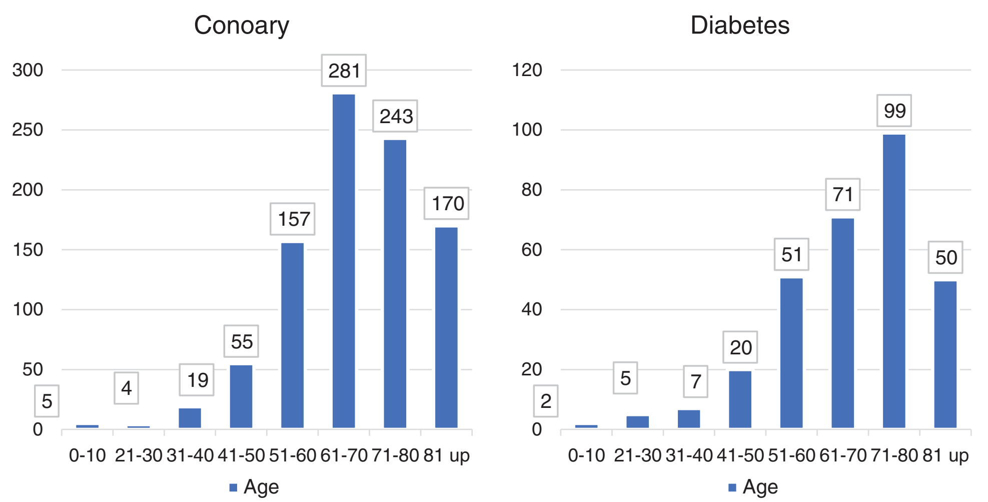


**Figure 3:** The distribution of age for coronary heart disease and diabetes

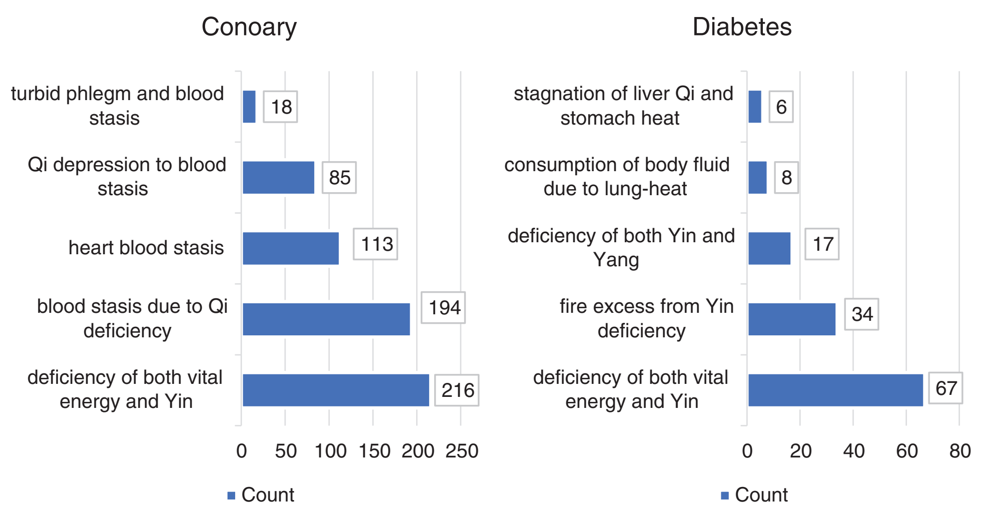


**Figure 4:** The distribution of syndromes for coronary heart disease and diabetes

**Table 3:** Experiment results on the clinical record dataset

| Disease | Count | Hit@1 | Hit@3 | Hit@5 | Meanrank |
|---|---|---|---|---|---|
| Coronary heart disease | 190 | 0.568 | 0.811 | 0.905 | 2.173 |
| Diabetes | 168 | 0.589 | 0.881 | 0.935 | 1.994 |
| Acyesis | 25 | 0.72 | 1.0 | 1.0 | 1.32 |
| Uterine bleeding | 15 | 0.733 | 1.0 | 1.0 | 1.333 |
| Menoxenia | 74 | 0.77 | 0.973 | 1.0 | 1.365 |
| Menorrhalgia | 17 | 0.882 | 1.0 | 1.0 | 1.118 |
| Menostasis | 20 | 0.9 | 1.0 | 1.0 | 1.1 |
| Menopausal syndrome | 10 | 0.9 | 1.0 | 1.0 | 1.1 |
| Mean of all | 519 | 0.708 | 0.958 | 0.980 | 1.438 |

In the real-world experiment, we let the doctors treat the diagnosis result as correct if one of the top three syndromes is consistent with their diagnosis. Otherwise, the result is considered wrong. In this experiment, there are 934 cases of coronary heart disease and 314 cases of diabetes. Tab. 4 displays the results of this evaluation. We can observe that the ratio of correct diagnoses is very high. Therefore, the proposed method has good performance in clinical practice.

**Table 4:** Experiment results in clinical practice

| Disease | Count | Correct (%) | Incorrect (%) |
|---|---|---|---|
| Coronary heart disease | 934 | 97.86 | 2.14 |
| Diabetes | 314 | 83.91 | 16.09 |

## 4 Discussion

Unlike most previous research that treats automated diagnosis as a supervised task, our method does not rely on a large annotation dataset for training with machine learning. We utilize the TCM knowledge graph and develop an unsupervised automated diagnosis method to achieve syndrome differentiation. Compared with other reported work, our method is robust and explainable under the guide of the knowledge graph. Our method's performance indicates its effectiveness in clinical practice. Moreover, our method could easily be generalized to other diseases.

However, the proposed method has limitations. First, it requires a high-quality knowledge graph. Thus, a stronger knowledge graph could improve its performance. Moreover, this method still relies on the prior knowledge of experts to a certain degree. Thus, we must consider introducing supervised learning. Lastly, additional clinical data are needed in future work.

## 5 Conclusion

Automated diagnosis is an essential and vital task. For TCM, syndrome differentiation is an important part of the diagnostic process. We propose an automated diagnosis method that simulates syndrome differentiation through deductive reasoning on a knowledge graph. We evaluate the method using a clinical record dataset and assess its application to clinical practice. The preliminary results suggest that the method can support diagnosis. It could help TCM physicians, especially primary physicians in rural areas, make clinical decisions. This will solve the imbalance of the medicine resource problem in China and lead to social and economic benefits.

**Acknowledgement:** We thank the anonymous reviewers for their helpful comments. Thanks are also due for the TCM knowledge support from Yingjie Shi and for the data processing by Hu Tao and Jia Li. We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

**Funding Statement:** This work is supported by the National Key Research and Development Program of China under Grant 2017YFB1002304 and the China Scholarship Council under Grant 201906465021.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

[1] S. K. Pal, "Complementary and alternative medicine: An overview," *Current Science*, vol. 82, no. 5, pp. 518–524, 2002.

[2] P. M. Barnes, E. Powell-Griner, K. McFann and R. L. Nahin, "Complementary and alternative medicine use among adults: United States," *2002 Seminars in Integrative Medicine*, vol. 2, no. 2, pp. 54–71, 2004.

[3] J. Qiu, "Traditional medicine: A culture in the balance," *Nature*, vol. 448, no. 7150, pp. 126–128, 2007.

[4] Y. N. Xiang, Z. M. Guo, P. F. Zhu, J. Chen and Y. Y. Huang, "Traditional Chinese medicine as a cancer treatment: Modern perspectives of ancient but advanced science," *Cancer Medicine*, vol. 8, no. 5, pp. 1958–1975, 2019.

[5] Ministry of Health, P.R. China, *China's Health Statistics Yearbook 2019*, 1st ed., Beijing, China: Peking Union Medical College Press, 2019.

[6] M. Jiang, C. Lu, C. Zhang, J. Yang, Y. Tan *et al.,* "Syndrome differentiation in modern research of traditional Chinese medicine," *Journal of Ethnopharmacology*, vol. 140, no. 3, pp. 634–642, 2012.

[7] D. L. Yu, L. Zhang, Y. G. Wang and Q. M. Zhang, "The diagnostic evidences of syndrome elements are symptom characteristics," *Chinese Journal of Basic Medicine in Traditional Chinese Medicine*, vol. 20, no. 12, pp. 1024–1025, 2014.

[8] H. Y. Liang, B. Y. Tsui, H. Ni, C. C. S. Valentim, S. L. Baxter *et al.,* "Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence," *Nature Medicine*, vol. 25, no. 3, pp. 433–438, 2019.

[9] Y. Juhn and H. F. Liu, "Artificial intelligence approaches using natural language processing to advance EHR-based clinical research," *Journal of Allergy and Clinical Immunology*, vol. 145, no. 2, pp. 463–469, 2020.

[10] J. Y. Hu, A. Perer and F. Wang, "Data driven analytics for personalized healthcare," in *Healthcare Information Management Systems*, 4th ed., vol. 1. Switzerland: Springer Cham Press, pp. 529–554, 2016.

[11] Y. Huang, Y. Gao and B. Ma, "Review of data mining methods frequently used in TCM syndrome study," *Acta Chinese Medicine and Pharmacology*, vol. 38, no. 3, pp. 6–10, 2010.

[12] J. Poon and S. K. Poon, *Data Analytics for Traditional Chinese Medicine Research*, 1st ed., vol. 1. Switzerland: Springer International Press, 2014.

[13] Y. Q. Wang, Z. H. Yu, Y. G. Jiang, Y. C. Liu, L. Chen *et al.,* "A framework and its empirical study of automatic diagnosis of traditional Chinese medicine utilizing raw free-text clinical records," *Journal of Biomedical Informatics*, vol. 45, no. 2, pp. 210–223, 2012.

[14] Q. Xu, W. J. Tang, F. Teng, W. Peng, Y. F. Zhang *et al.,* "Intelligent syndrome differentiation of traditional chinese medicine by ANN: A case study of chronic obstructive pulmonary disease," *IEEE Access*, vol. 7, pp. 76167–76175, 2019.

[15] Z. Q. Liu, H. Y. He, S. X. Yan, Y. Wang, T. Yang *et al.,* "End-to-end models to imitate traditional Chinese medicine syndrome differentiation in lung cancer diagnosis: Model development and validation," *JMIR Medical Informatics*, vol. 8, no. 6, pp. e17821, 2020.

[16] H. Zhang, W. D. Ni, J. Li and J. J. Zhang, "Artificial intelligence-based traditional Chinese medicine assistive diagnostic system: Validation study," *JMIR Medical Informatics*, vol. 8, no. 6, pp. e17608, 2020.

[17] J. Lehmann, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas *et al.,* "DBpedia—A large-scale, multilingual knowledge base extracted from wikipedia," *Semantic Web*, vol. 6, no. 2, pp. 167–195, 2015.

[18] H. Lian, Z. M. Qin, T. K. He and B. Luo, "Knowledge graph construction based on judicial data with social media," in *Proc. Web Information Systems and Applications Conf.*, Liuzhou, Guangxi Province, China, pp. 225–227, 2017.

[19] Y. H. Xie, C. Yan and D. Z. Zhang, "Personalized diagnostic modal discovery of traditional Chinese medicine knowledge graph," in *Proc. Int. Conf. on Natural Computation, Fuzzy Systems and Knowledge Discovery*, Huangshan, Anhui Province, China, pp. 1096–1103, 2018.

[20] T. Yu, J. H. Li, Q. Yu, Y. Tian, X. F. Shun *et al.,* "Knowledge graph for TCM health preservation: Design, construction, and applications," *Artificial Intelligence in Medicine*, vol. 77, pp. 48–52, 2017.

[21] Z. Q. Zheng, Y. G. Liu, Y. Zhang and C. B. Wen, "TCMKG: A deep learning based traditional Chinese medicine knowledge graph platform," in *Proc. IEEE Int. Conf. on Knowledge Graph*, Nanjing, Jiangsu Province, pp. 560–564, 2020.

[22] D. Z. Zhang, Y. H. Xie, M. Li and C. Shi, "Construction of knowledge graph of traditional Chinese medicine based on the ontology," *Information Engineering*, vol. 3, no. 1, pp. 35–42, 2017.

[23] Y. Xie, L. Hu, X. Chen, J. Feng and D. Zhang, "Auxiliary diagnosis based on the knowledge graph of tcm syndrome," *Computers, Materials & Continua*, vol. 65, no. 1, pp. 481–494, 2020.

## Appendix A. The Chinese–English translations of syndromes

| Chinese | English |
|---|---|
| 肺热津伤证 | Consumption of body fluid due to lung-heat |
| 肝气郁滞证 | Stagnation of liver-Qi |
| 肝肾不足证 | Deficiency of liver and kidney |
| 肝肾亏损证 | Deficiency of the liver and kidney |
| 肝肾阴虚证 | Liver-kidney Yin deficiency |
| 肝胃郁热证 | Stagnation of liver Qi and stomach heat |
| 肝血亏虚证 | Deficiency of liver-blood |
| 肝郁血热证 | Stagnation of liver Qi and blood stasis |
| 寒凝心脉证 | Cold stagnation in heart vessel |
| 寒凝血瘀证 | Cold stagnation and blood stasis |
| 寒湿凝滞证 | Cold-dampness to freeze and sluggish |
| 精血亏虚证 | Blood and essence asthenia |
| 脾气虚证 | Deficiency of spleen-Qi |
| 脾肾阳虚证 | Insufficiency of both the spleen and the kidney |
| 脾虚痰湿证 | Deficiency of the spleen causing stagnation of |
| 脾虚证 | splenasthenic syndrome |
| 气虚血瘀证 | Blood stasis due to Qi deficiency |
| 气血不足证 | Insufficiency of vital energy and blood |
| 气血亏虚证 | Qi-blood deficiency |
| 气阴两虚证 | Deficiency of both vital energy and Yin |
| 气滞心胸证 | Qi stagnation in heart and chest |
| 气滞血瘀证 | Qi depression to blood stasis |
| 热毒血瘀证 | Toxic-heat and blood stasis |
| 肾气虚证 | Kidney Qi deficiency |
| 肾虚肝郁证 | Kidney deficiency and liver-Qi stagnation |
| 肾虚痰湿证 | Deficiency of the kidney causing stagnation of phlegm |
| 肾虚血亏证 | Deficiency of the kidney and blood |
| 肾虚血瘀证 | Blood stasis and kidney deficiency |
| 肾阳虚证 | Kidney-Yang deficiency |
| 肾阴虚证 | Kidney-Yin deficiency |
| 肾阴阳俱虚证 | Kidney Yin and Yang deficiency |
| 湿热瘀结证 | Damp-heat gluing |
| 湿热瘀阻证 | Dampness and heat stasis |
| 实热证 | Excess-heat syndrome |
| 痰热互结证 | Combination of phlegm and heat |
| 痰湿内阻证 | Retention of phlegm-dampness in the interior |
| 痰湿阻滞证 | Phlegm stagnation |
| 痰瘀互结证 | Intermin-gled phlegm and blood stasis |
| 痰浊闭阻证 | Turbid phlegm blocking |
| 痰阻血瘀证 | Turbid phlegm and blood stasis |
| 胃热津亏证 | Insufficiency of both the spleen and the kidney |
| 胃燥津伤证 | Fluid injury due to stomach dryness |
| 心气虚弱证 | Deficiency of heart Qi |
| 心肾不交证 | Failure of the heart and kidney integrating |

| Chinese | English |
| --- | --- |
| 心肾阳虚证 | Heart-kidney Yang deficiency |
| 心肾阴虚证 | Heart-kidney Yin deficiency |
| 心血瘀阻证 | Heart blood stasis |
| 虚热证 | Asthenia pyrosyndrome |
| 血寒凝滞证 | Pathogenic cold coagulating with blood |
| 阳盛血热证 | Yang deficiency and blood heat |
| 阴虚火旺证 | Fire excess from Yin deficiency |
| 阴虚内热证 | Yin asthenia generating intrinsic heat |
| 阴虚血热证 | Yin deficiency and blood heat |
| 阴虚血燥证 | Yin-deficiency and blood-dryness |
| 阴虚阳浮证 | Floating of Yang due to Yin deficiency |
| 阴阳两虚证 | Deficiency of both Yin and Yang |
| 瘀滞胞宫证 | Stagnation in uterus |
| 正阳虚脱证 | Yang prostration |