Neural network analysis of Chinese herbal medicine prescriptions for patients with colorectal cancer

Yu-Chuan Lin[a], Wei-Te Huang[a], Shi-Chen Ou[a], Hao-Hsiu Hung[a], Wie-Zen Cheng[a], Sheng-Shing Lin[a], Hung-Jen Lin[a], Sheng-Teng Huang[a,b,c,d,e,f,*]

[a] Department of Chinese Medicine, China Medical University Hospital, Taichung, Taiwan
[b] School of Chinese Medicine, China Medical University, Taichung, Taiwan
[c] Research Center for Traditional Chinese Medicine, Department of Medical Research, China Medical University Hospital, Taichung, Taiwan
[d] Chinese Medicine Research Center, China Medical University, Taichung, Taiwan
[e] Research Center for Chinese Herbal Medicine, China Medical University, Taichung, Taiwan
[f] Tainan Municipal An-Nan Hospital, China Medical University, Taichung, Taiwan

ARTICLE INFO

ABSTRACT

Traditional Chinese Medicine (TCM) is an experiential form of medicine with a history dating back thousands of years. The present study aimed to utilize neural network analysis to examine specific prescriptions for colorectal cancer (CRC) in clinical practice to arrive at the most effective prescription strategy. The study analyzed the data of 261 CRC cases recruited from a total of 141,962 cases of renowned veteran TCM doctors collected from datasets of both the DeepMedic software and TCM cancer treatment books. The DeepMedic software was applied to normalize the symptoms/signs and Chinese herbal medicine (CHM) prescriptions using standardized terminologies. Over 20 percent of CRC patients demonstrated symptoms of poor appetite, fatigue, loose stool, and abdominal pain. By analyzing the prescription patterns of CHM, we found that Atractylodes macrocephala (Bai-zhu) and Poria (Fu-ling) were the most commonly prescribed single herbs identified through analysis of medical records, and supported by the neural network analysis; although there was a slight difference in the sequential order. The study revealed an 81.9% degree of similarity of CHM prescriptions between the medical records and the neural network suggestions. The patterns of nourishing Qi and eliminating dampness were the most common goals of clinical prescriptions, which corresponds with treatments of CRC patients in clinical practice. This is the first study to employ machine learning, specifically neural network analytics to support TCM clinical diagnoses and prescriptions. The DeepMedic software may be used to deliver accurate TCM diagnoses and suggest prescriptions to treat CRC.

1. Introduction

Traditional Chinese medicine (TCM) is an ancient health care system that has been trusted to manage the health care needs of the wider Asian population for thousands of years. As such, it benefits from the accrued knowledge of TCM practitioners over this timespan, and has been refined by billions of clinical experiences. In terms of the approach to treatment of ailments, TCM primarily focuses on harmonizing the Yin-Yang balance.[1] Though its approach to disease treatment may not be as focused and aggressive as western medicine, TCM treatment modalities can make patients feel comfortable by promoting homeostasis. As such, it is commonly adopted at present as an adjunct modality to improve the quality of life in patients after receiving western medical treatment, and has been rediscovered as a popular treatment alternative in Asia over recent decades.[2]

However, traditional Chinese medicine arguably suffers from various drawbacks or weaknesses. One of the foremost of which is that TCM practitioners may have different prescriptions for individual patients, based on the practitioners' unique education and personal clinical experiences. Even renowned veteran TCM doctors will demonstrate differing treatments. The lack of standardized diagnosis and prescription behavior results in not only difficult or confusing learning environments for young doctors but also uncertainty or confusion for patients. Fortunately, this situation can be remedied through the

application of big data and machine learning, which are undergoing rapid development and innovation, and are currently being applied in various fields.

For the present study, we used data extraction techniques, including e-book data structuring, data mining, and knowledge discovery to gather the treatment experiences of renowned veteran TCM doctors.[3,4] Neural networks linking symptoms, patterns, syndrome elements, Chinese herbs and formulas were applied by utilizing natural language processing to translate data into a standardized language so the neural network could more easily learn from and analyze the data.[5,6] This software can transform the information found in medical records into standardized symptoms and output suggested diagnoses and prescriptions for treatment.[7] Furthermore, the cosine similarity was applied to evaluate the matching of these prescriptions between original medical records and the machine learning suggestions. The suggested prescriptions may be used to reach a consensus among current TCM clinical practitioners.

This is the first study to explore and compare TCM patterns by utilizing a machine learning neural network to simulate diagnosis and arrive at the most effective prescriptions. As the third most commonly occurring cancer in the world, this study focused on cases of CRC.[8,9] We aim to apply big data neural network analysis in TCM clinical practice to educate doctors by improving and standardizing the treatment of various diseases and ailments. The collection of ancient knowledge and subsequent analysis will not only help to prevent diagnostic error in clinic practice but also allow junior TCM doctors or researchers to follow standardized diagnosis and prescription patterns in their education and research.

2. Materials and methods

2.1. Data acquisitions

The CRC cases were recruited from 141,962 cases of renowned veteran TCM doctors collected from datasets of both the DeepMedic (DM) software and TCM cancer treatment books, as shown in Supplementary Table 1. The key word "intestines" was used to search the data of western medical terms, revealing 2,277 cases. We subsequently excluded records of non-colon cancer, records without use of Chinese herbal medicine, and unanalyzable records. Thus, 261 CRC cases were recruited for further analysis. The flowchart of our data acquisition process is shown in Fig. 1.

2.2. Standardized terminologies of traditional Chinese medicine

The standardization tool designed by DeepMedic was applied to normalize the symptoms/signs and the prescriptions of Chinese herbal medicine in the 261 cases. The details of the related methods are demonstrated in the Supplementary materials. The symptoms, signs, syndrome elements, and CHM prescriptions of the CRC cases were analyzed and then compared between the medical records and the neural network output.

2.3. DeepMedic neural network analysis

The DeepMedic software is a tool used to standardize and analyze the terminologies of traditional Chinese medicine and to suggest the most effective prescriptions. The framework of DeepMedic is depicted in the Supplementary materials. The website accessing the demo version of DeepMedic may be found at: http://bigdata-demo.deepmedic. cn/.

The main purpose of the standardization process is to unify the polysemous or synonymous vocabulary for TCM verification, validation, and accreditation to facilitate the machine learning analysis and output.

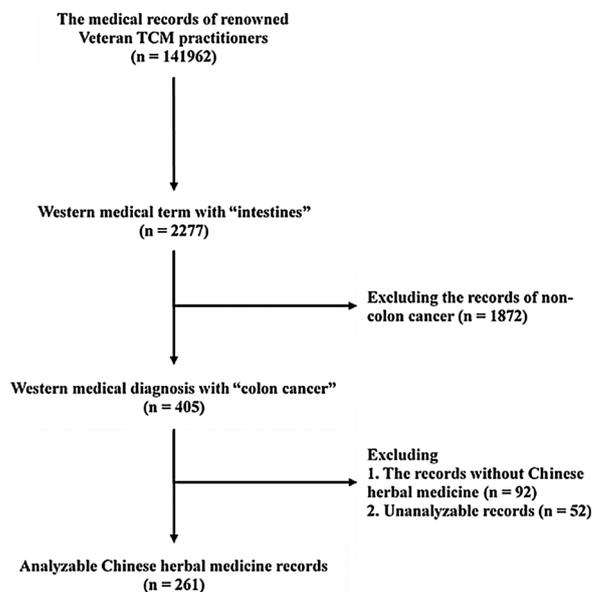The standardization of TCM symptoms is performed via the neural

Fig. 1. Flow chart of study cases extracted from renowned veteran TCM practitioners.

network, facilitated by modifying symptom vocabulary to match the over 20,000 symptom terminologies of the thesaurus within the DeepMedic software. The standardization of TCM nomenclature is based on the authoritative Chinese medicine pharmacology dictionary. The standard herb name and alias comparison tables are also applied to standardize the herb names appearing in the cases. The standardization process of syndrome elements, TCM patterns, Western medicine diagnoses, and treatment modalities were unified by using their respective standard nomenclatures.

In the process of standardization, DeepMedic will carry out separate analyses for both the symptoms and medications of each case. It also exerts intelligent pattern identification, analysis of prescription efficacy, and matching of percentages for formulas and prescription patterns. The DeepMedic software selects cases with pattern identification, disassembles those patterns into several codes, and labels the standard symptoms or standard medication used in the case. Thus, for each case to be analyzed as input, DeepMedic can quickly identify the higher-weighted code group composed of different types of patterns, resulting in pattern identification. Subsequently, DM will identify the corresponding code group associated with herbs and formulas and output a recommended prescription. Moreover, DeepMedic can identify the higher weighting code group from the original data of each case to trace back the accuracy of the prescription and the logical thought process of clinical practitioners. Additionally, the approximation between the symptoms and medication can be calculated by cosine vector formula, to compare the matching percentages between formulas and patterns.

Cosine vector formula is the inner product of two high dimension vectors divided by their distances which represent the formula and pattern conclusion. The dimensions are the codes that have TCM meanings: $COS(f, p) = (f \cdot p) / |f||p|$

The code dimensions are semi-correlated with the TCM meaning and the output of the neural network is between 0 and 1. Thus, simply using COS function can give us a straightforward impression of how to match the formulas and patterns without involving complexity.

The weightings of each code based on different symptoms are calculated by a back-forward propagation of neural network consisting of several hidden layers. The reason for not choosing deeper hidden layers is because the relation between the codes and the symptoms is more direct in the TCM field, therefore this prevents overfitting. The weightings of each pattern are based on different patterns calculated by using the well-known heuristic equation, Term-Frequency-Inverse

Document Frequency (TF-IDF) with some modifications. The consideration of TF-IDF is that the TCM diagnosis relies on the discrimination of natural signs. Thus, we smooth the calculating results of TF-IDF in order to know the strength of one specific symptom to identify the pattern.

TF = (the frequencies of symptom A in code B / code)

$$Term\ frequency = f_{t,d} / \sum_{t \in d} f_{t',d}$$

$$Inverse\ document\ frequency\ smooth = \log(1 + N/n_t)$$

2.4. Data analysis

We calculated the frequency of occurrence for the top ten standard major symptoms, as well as for the top five pulses and tongue diagnoses of the CRC cases. We further compared the ten most common herbal prescriptions between the medical records and the neural network output suggestion. The Pearson correlation coefficient (PCC) was used to indicate the similarities between these two comparisons. The NodeXL software (https://nodexl.com/) was applied to interpret the associations of both standard symptoms and medication.

3. Results

3.1. The main symptoms and TCM clinical diagnoses in patients with CRC

TCM clinical data contains distinctive information, including TCM symptoms and signs used to diagnose patterns and arrive at a prescription for treatment. In this study, the top 10 main symptoms are shown in Table 1. Over 20 percent of patients with CRC demonstrated symptoms of poor appetite (n = 115, 44.1%), fatigue (n = 102, 39.1%), loose stool (n = 61, 23.4%), and abdominal pain (n = 55, 21.1%). While less than 20 percent of reported symptoms included insomnia (17.6%), abdominal distension (14.6%), dry mouth and nausea (10.7%), emaciation (10.0%), and hematochezia (9.6%).

According to TCM theory, the tongue and pulse examinations are of great importance to the diagnostic procedure. The diagnosed patterns illuminate what treatment options are advisable, and what symptoms need to be addressed. Thus, we further analyzed the pulse and tongue signs of the CRC patients. A fine pulse was the most common pulse diagnosis (61.7%), followed by string-like pulse (35.6%), deep pulse (27.2%), slippery pulse (17.2%), and weak pulse (8.4%), as shown in Table 2. Furthermore, the top 5 tongue diagnoses were red (18.0%), dark red (17.2%), light red (11.5%), ecchymosis (7.7%), and teeth marks (4.6%). As for the tongue coating, the sequential order was thin (41.8%), white (41.0%), yellow (33.7%), greasy (26.4%), and thin with white (20.3%), as shown in Table 2.

An analysis was applied in this study in order to link the dominance and association between symptoms and signs present in the TCM diagnostic patterns, as shown in Fig. 2. The fine pulse was dominant in

CRC patients highly associated with the other four major symptoms and signs, including thin coating (5), poor appetite (4.93), white coating (4.8), and fatigue (4.4), indicating a Qi deficiency and dampness.

3.2. Comparison of prescriptions of Chinese herbal medicine for CRC patients

Herbal medicine prescription is based on the symptoms and signs identified during the TCM diagnostic process, and is applied to reach internal homeostasis. However, different practitioners could apply unique pattern diagnoses, and therefore arrive at differing prescriptions for treatment. Thus, we compared the differences of prescriptions between renowned veteran TCM practitioners and the suggested prescriptions of the neural network. We found that *Atractylodes macrocephala* (Bai-zhu, 52.3%) and *Poria* (Fu-ling, 52.3%) were the most commonly prescribed single herbs according to medical records, followed by *Radix Astragali* (Huang-Qi, 43.4%), *Radix codonopisis* (Dang-shen, 26.3%), and *Semen Coicis* (Yi-yi-ren, 23.5%); meanwhile, there were minor differences in the sequence of prescriptions suggested by the neural network, including *Atractylodes macrocephala* (Bai-zhu, 40.6%) and *Poria* (Fu-ling, 35.9%), *Radix Glycyrrhizae Preparta* (Gan-cao, 32.7%), *Radix codonopisis* (Dang-shen, 29.2%) and *Radix Bupleuri* (Chai-hu, 26.2). These results demonstrated a high degree of similarity between the TCM doctors' prescriptions and those of the neural network output (Pearson correlation coefficient, PCC = 81.9%) (see Table 3). Most significantly, the network analysis demonstrated a high correlation of *Atractylodes macrocephala*, *Radix Astragali*, *Poria*, *Radix codonopisis* and *Radix Glycyrrhizae* between medical records and the neural network output, indicating the importance of nourishing Qi and eliminating dampness for patients with CRC in clinical practice (see Fig. 3).

4. Discussion

TCM is an ancient form of health care based on the accrued wisdom and experience of practitioners. While effective clinical experiences and prescriptions were recorded for the benefit of future generations, the sheer amount of data gathered over thousands of years is onerous. However, due to technological advances, this ancient wisdom is presently being transformed into a digital format by optical character recognition (OCR) by the DeepMedic team. Thus, this study attempts to integrate the information collected from renowned veteran TCM clinical practitioners and utilize neural network analysis to identify the most effective Chinese herbal medicine treatments for patients with CRC. To the best of our knowledge, this is the first study to use a machine learning, specifically neural network analysis, to arrive at the most effective prescriptions for clinical practitioners.

At present, the most accurate methods for the diagnosis of CRC are by sigmoidoscopy and colonoscopy examinations.[10] According to reports, bloody stool and change of bowel movement habits are the most common symptoms occurring in patients with CRC.[11,12] Thus, consideration of patients' symptoms and signs may reveal important clues to be used for the early diagnosis of CRC. In the present study, we reveal that the top ten symptoms include six gastrointestinal issues, and the frequency of hematochezia is 9.6%. Additionally, treatments such as chemotherapy, radiotherapy, and/or surgery in CRC patients have been reported to result in physical discomfort, including dyspepsia, sleep disturbance, and GI upsets.[13] This study found that poor appetite is the chief symptom, while psychological and physical illnesses such as fatigue and insomnia are also common. These symptoms may make patients feel a general sense of weakness and discomfort, and may consequently restrict further western medical treatments.

In addition to the aforementioned, pulse and tongue diagnoses are essential tools in the assessment of the homeostasis of Ying-Yang and the five major organ systems, according to TCM theory. A fine pulse, the most common pulse pattern identified in our study, is associated with Qi and blood deficiencies, indicating a chronic illness. A red or dark red

Table 1
Top ten of the chief symptoms of patients with colorectal cancer.

Symptoms	Frequency (%)
Poor appetite	115(44.1)
Fatigue	102(39.1)
Loose stool	61(23.4)
Abdominal pain	55(21.1)
Insomnia	46(17.6)
Abdominal distension	38(14.6)
Dry mouth	28(10.7)
Nausea	28(10.7)
Emaciation	26(10.0)
Hematochezia	25(9.6)

Table 2
Top five TCM clinical diagnoses of pulse and tongue in patients with colorectal cancer.

Pulse diagnosis		Tongue diagnosis			
Type	Frequency (%)	body	Frequency (%)	coating	Frequency. (%)
Fine pulse	161(61.7)	Red	47(18.0)	Thin coating	109(41.8)
String-like pulse	93(35.6)	Dark red	45(17.2)	White coating	107(41.0)
Deep pulse	71(27.2)	Light red	30(11.5)	Yellow coating	88(33.7)
Slippery pulse	45(17.2)	Ecchymosis	20(7.7)	Greasy coating	69(26.4)
Weak pulse	22(8.4)	Teeth Marks	12(4.6)	Thin with white coating	53(20.3)

tongue coloration combined with a white or yellow coating represents heat repletion or Ying vacuity associated with dampness.[14,15] Furthermore, previous studies have reported that CRC patients' tongue coatings demonstrated obvious thickening in comparison with healthy people. Interestingly, a purple tongue, equivalent to a dark red tongue, accompanied ecchymosis in approximately 24.9% of our study population, and appears frequently in most cancer patients to indicate the pattern of blood stasis.[16,17]

Cancer is a complicated systemic disease often affecting multiple organs. According to TCM theory, tumors initially act as a kind of toxic heat to exhaust the human Qi and blood, eventually leading to Qi stagnation and blood stasis. As such, a majority of TCM practitioners prescribe herbs to nourish the Qi and blood, clear heat, and to remove blood stasis, in accordance with the differing diagnostic patterns associated with cancer patients. Previous studies have reported the most frequently prescribed formulas in patients with CRC was Xiang-Sha-Liu-Jun-Zi-Tang, composed of *Vladimiria radix* (Mu-Xiang), *Villous amomum* (Sha-Ren), *Citri reticulatae* (Chen-Pi), *Pinellia ternate* (Ban-Xia), *Ginseng radix* (Ren-Shen), *Atractylodes macrocephala* (Bai-Zhu), *Poria* (Fu-Ling), *Glycyrrhizae preparta* (Gan-Cao), *Zingiberis rhizome* (Gan-Jiang), *Zizyphus fructus* (Da-Zao). In addition, the single herbs*Hedyotis diffusa*

(Bai-Hua-She-She-Cao), *Scutellaria barbata* (Ban-Zhi-Lian), *Ginseng Radix* (Ren-Shen) and *Radix Astragali* (Huang-Qi) are also commonly prescribed.[13,18] Most of these herbs can invigorate the spleen and stomach so as to ensure production of blood and Qi to improve quality of life and reduce side effects induced by western medical treatment.[19] This study identifies significant similarities of TCM prescriptions among previous studies, renowned veteran TCM doctors, and the neural network analysis output.

As for the specific anti-cancer effects of the abovementioned CHM prescriptions, *Atractylodes macrocephala* suppresses β-catenin/T-cell factor transcriptional activity to downregulate the nuclear level of β-catenin and inhibits proliferation of CRC cells.[20] *Poria* has been reported to potentiate immune response and inhibit tumor growth by anti-angiogenesis via downregulation of both NF-κB and NF-κB/Rel translocation.[21] *Radix Astragali* exerts anti-angiogenic effects by reducing VEGF ligand/receptor and anti-carcinogenesis through Akt/mTOR and COX-2 pathways to inhibit CRC cell growth both in vitro and in vivo.[22] Codonoposide 1c, isolated from *Radix codonopisis*, activates caspase cascade through the intrinsic pathway to modulate Bid, Bax, and Smac to induce apoptosis and inhibit cancer growth.[23] *Pericarpium Citri Reticulatae* induces apoptosis in human colon cancer cells through
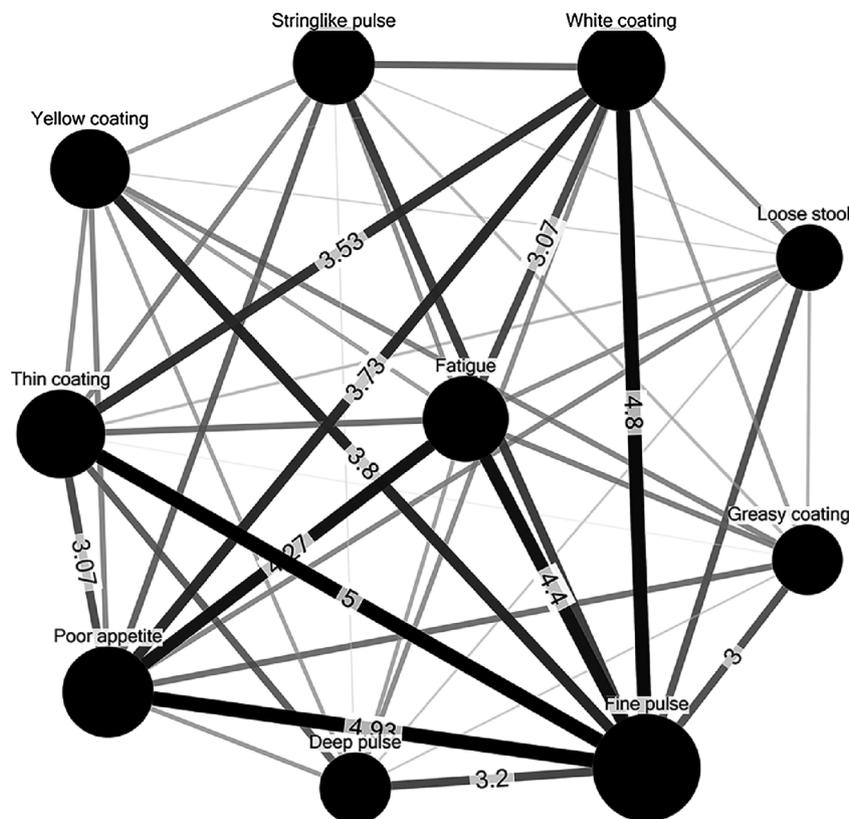
Fig. 2. Network of standard symptoms. Network analysis of the ten most frequent symptoms and signs for all patients with colorectal cancer. The spot size indicates the frequency of clinical observation, and the line width indicates the association between symptoms and signs.

Table 3

Top ten single Chinese herbs prescribed for colorectal cancer patients between medical record and neural network output.

Medical record latin name (Pin-Yin name)	Frequency (%)	Neural network output latin name (Pin-Yin name)	Frequency (%)
Atractylodes macrocephala (Bai-zhu)	147(52.3)	*Atractylodes macrocephala* (Bai-zhu)	114(40.6)
Poria (Fu-ling)	147(52.3)	*Poria* (Fu-ling)	101(35.9)
Radix Astragali (Huang-Qi)	122(43.4)	*Radix Glycyrrhizae Preparta* (Gan-cao)	92(32.7)
Radix codonopsis (Dang-shen)	74(26.3)	*Radix codonopsis* (Dang-shen)	82(29.2)
Semen Coicis (Yi-yi-ren)	66(23.5)	*Radix Bupleuri* (Chai-hu)	73(26.0)
Radix Glycyrrhizae Preparta (Gan-cao)	65(23.1)	*Radix Astragali* (Huang-Qi)	71(25.3)
Pericarpium Citri Reticulatae (Chen-pi)	63(22.4)	*Radix Paeoniae Lactiflorae* (Bai-shao-yao)	63(22.4)
Pseudostellariae Radix (Tai-zi-shen)	59(21.0)	*Radix Angelicae sinensis* (Dang-gui)	61(21.7)
Hedyotis diffusa Willd (Bai-hua-she-she-cao)	54(19.2)	*Pericarpium Citri Reticulatae* (Chen-pi)	51(18.1)
Rhizoma Pinelliae (Ban-xia)	51(18.1)	*Radix Salviae Miltiorrhizae* (Dan-shen)	43(15.3)

PCC[a] = 81.9% indicated the similarity between medical record and neural network output.

[a]PCC Pearson correlation coefficient.

Bax-related caspase-3 activation.[24] While Lin et al. has reported that *Hedyotis diffusa* suppresses CRC cell growth via inducing apoptosis, and inhibiting angiogenesis through the VEGF pathway as well as the cancer-promoting inflammatory environment through IL-6/STAT3 signaling modulation.[25,26] Tanshinone I and II, major extracts from *Radix Salviae Miltiorrhizae*, induce CRC apoptosis and G0/G1 cell growth arrest by p53 and survivin-mediated signaling pathways.[27,28] In addition, *Semen Coicis*, a common component in cereal products, suppresses early events in colon carcinogenesis by reducing the number of preneoplastic aberrant crypt foci (ACF) associated with the decrease of COX-2 protein expression.[29,30] Moreover, DangguiBuxue Tang, composed of *Radix Angelicae sinensis* and *Radix Astragali*, induces autophagy-associated cell death in CRC cells as a chemotherapy or radiotherapy sensitizer in CRC treatment.[31] The possible therapeutic and pharmacological effects of these ten single herbs in CRC management are illustrated in Supplementary Table 2.[32–78]

In the future, neural network analysis to examine colorectal cancer (CRC) cases should also include consideration of the multi-scale and dynamic behaviors of CRC. We know that inflammation has a close relationship with cancer, including responses to DNA damage, gene mutation, cell cycle behavior, population dynamics, and metabolism-immune balance.[79] Therefore, apart from TCM prescription and pattern analysis, modern medicine laboratory data, molecular biology, and data from other fields should be included in the analysis module in the

future.

There are several inherent limitations to this study which must be addressed. First, the software used herein involves a neural network and natural language processing which may neglect some symptoms and signs potentially used for more detailed differential diagnosis. Second, data are insufficient to distinguish the specific treatments effective at relieving side effects or inhibiting cancer growth. Finally, the data shown in this study only present diagnosis and/or treatment from Chinese medical literature as compared with the neural network analysis output; however, the similarities in CRC treatment prescriptions between TCM practitioners and the neural network output require further validation based on TCM theory, and necessitate further clinical observation and trials to verify effectiveness.

5. Conclusion

This is the first study to utilize neural network analytics to suggest clinical prescriptions through the analysis of big data resources. The DeepMedic software applied in the present study identified herbs with the purpose of nourishing Qi and eliminating dampness for its suggested prescriptions, accurately corresponding to TCM prescriptions for the treatment of CRC in clinical practice. The successful application of this technology in TCM will improve the effectiveness of disease diagnosis and treatment, and may be applied to enhance education and
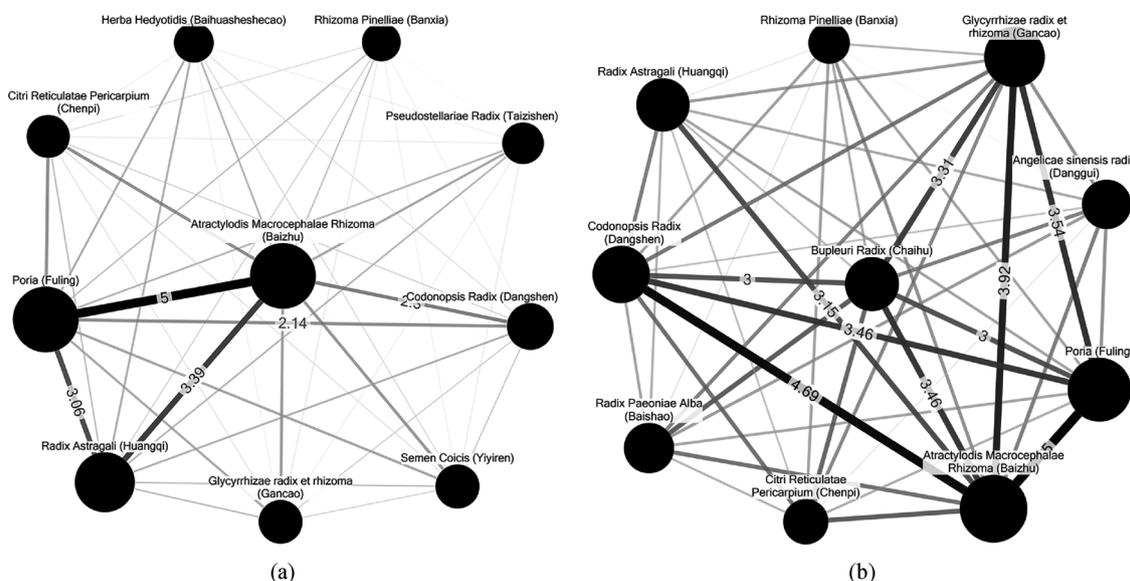
Fig. 3. Network of medicine. Network analysis of the ten most frequent herbs prescribed by renowned veteran TCM practitioners (a) and neural network (b) in patients with colorectal cancer. The spot size indicates the frequency of Chinese herbal product prescription, and the line width indicates the association between two Chinese herbs.

research in the future.

Conflicts of interest

The authors declare that they have no competing interests.

Authors' contributions

YCL wrote the draft and interpreted the data. WDH, SCO, HHH, WZC, SSL collected, assembled and analyzed the data. HJL provided study materials and administrative support. STH designed, conceived the study and amended the manuscript. YCL, WDH, SCO, HHH, WZC, SSL HJL, STH approved the final manuscript.

Acknowledgements

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ctim.2018.12.001.

References

1. Tang JL, Liu BY, Ma KW. Traditional Chinese medicine. *Lancet.* 2008;372(December (9654)):1938–1940https://www.ncbi.nlm.nih.gov/pubmed/18930523.

2. A. Burton, T.F. M. Smith, Q. Zhang, X. Zhang, T. Boerma, W. Lerberghe, Q. Zhang, E. Asio, Y. Maruyama. WHO traditional medicine strategy: 2014-2023. WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland: World Health Organization; 2013.http://www.who.int/medicines/publications/traditional/trm_strategy14_23/en/.

3. Zhou X, Peng Y, Liu B. Text mining for traditional Chinese medical knowledge discovery: a survey. *J Biomed Inform.* 2010;43(August (4)):650–660https://www.ncbi.nlm.nih.gov/pubmed/20074663.

4. Feng Y, Wu Z, Zhou X, Zhou Z, Fan W. Knowledge discovery in traditional Chinese medicine: State of the art and perspectives. *Artif Intell Med.* 2006;38(November (3)):219–236https://www.ncbi.nlm.nih.gov/pubmed/16930966.

5. Zhao C, Li GZ, Wang C, Niu J. Advances in patient classification for traditional chinese medicine: a machine learning perspective. *Evid Based Complement Alternat Med.* 2015;376716 2015 https://www.ncbi.nlm.nih.gov/pubmed/26246834.

6. Wang Y, Yu Z, Jiang Y, Xu K, Chen X. Automatic symptom name normalization in clinical records of traditional Chinese medicine. *BMC Bioinform.* 2010;20(January (11)):40https://www.ncbi.nlm.nih.gov/pubmed/20089162.

7. You M, Chen C, Li GZ, et al. ISMAC: an intelligent system for customized clinical case management and analysis. *Sci World J.* 2015;2015:473168https://www.ncbi.nlm.nih.gov/pubmed/26495425.

8. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer.* 2015;136(March (5)):E359–E386 1 https://www.ncbi.nlm.nih.gov/pubmed/25220842.

9. Salehiniya H, Pouyesh V, Tarazoj AA, et al. Colorectal cancer in the world: Incidence, mortality and risk factors. *Biomed Res Ther.* 2017;4(10):1656.

10. Swiderska M, Choromanska B, Dabrowska E, et al. The diagnostics of colorectal cancer. *Contemp Oncol (Pozn).* 2014;18(1):1–6https://www.ncbi.nlm.nih.gov/pubmed/24876814.

11. Jensen LF, Hvidberg L, Pedersen AF, Vedsted P. Symptom attributions in patients with colorectal cancer. *BMC Fam Pract.* 2015;3(September(16)):115https://www.ncbi.nlm.nih.gov/pubmed/26335940.

12. John SK, George S, Primrose JN, Fozard JB. Symptoms and signs in patients with colorectal cancer. *Colorectal Dis.* 2011;13(January (1)):17–25https://www.ncbi.nlm.nih.gov/pubmed/20105201.

13. Chao TH, Fu PK, Chang CH, et al. Prescription patterns of Chinese herbal products for post-surgery colon cancer patients in Taiwan. *J Ethnopharmacol.* 2014;155(August (1)):702–708.

14. Bilton K, Hammer L, Zaslawski C. Contemporary Chinese pulse diagnosis: A modern interpretation of an ancient and traditional method. *J Acupunct Meridian Stud.* 2013(October (5)):227–233https://www.ncbi.nlm.nih.gov/pubmed/24139459.

15. Lo LC, Chen CY, Chiang JY, Cheng TL, Lin HJ, Chang HH. Tongue diagnosis of traditional Chinese medicine for rheumatoid arthritis. *Afr J Tradit Complement Altern Med.* 2013;10(5):360–369https://www.ncbi.nlm.nih.gov/pubmed/24311851.

16. Han S, Chen Y, Hu J, Ji Z. Tongue images and tongue coating microbiome in patients with colorectal cancer. *Microb Pathog.* 2014;77(December):1–6https://www.ncbi.nlm.nih.gov/pubmed/25281933.

17. Han S, Yang X, Qi Q, et al. Potential screening and early diagnosis method for cancer: Tongue diagnosis. *Int J Oncol.* 2016;48(June (6)):2257–2264https://www.ncbi.nlm.nih.gov/pubmed/27035407.

18. Wu CT, Tsai YT, Lai JN. Demographic and medication characteristics of traditional Chinese medicine users among colorectal cancer survivors: a nationwide database study in Taiwan. *J Tradit Complement Med.* 2017;7(April (2)):188–194.

19. Nie J, Zhao C, Deng LI, et al. Efficacy of traditional Chinese medicine in treating cancer. *Biomed Rep.* 2016;4(January (1)):3–14https://www.ncbi.nlm.nih.gov/pubmed/26870326.

20. Shim AR, Dong GZ, Lee HJ, Ryu JH. Atractylochromene Is a Repressor of Wnt/beta-Catenin Signaling in Colon Cancer Cells. *Biomol Ther (Seoul).* 2015;23(January (1)):26–30https://www.ncbi.nlm.nih.gov/pubmed/25593640.

21. Rios JL. Chemical constituents and pharmacological properties of Poria cocos. *Planta Med.* 2011;77(May (7)):681–691https://www.ncbi.nlm.nih.gov/pubmed/21347995.

22. Law PC, Auyeung KK, Chan LY, Ko JK. Astragalus saponins downregulate vascular endothelial growth factor under cobalt chloride-stimulated hypoxia in colon cancer cells. *BMC Complement Altern Med.* 2012;19(September (12)):160https://www.ncbi.nlm.nih.gov/pubmed/22992293.

23. Lee KW, Jung HJ, Park HJ, Kim DG, Lee JY, Lee KT. Beta-D-xylopyranosyl-(1–&3)-beta-D-glucuronopyranosyl echinocystic acid isolated from the roots of Codonopsis lanceolata induces caspase-dependent apoptosis in human acute promyelocytic leukemia HL-60 cells. *Biol Pharm Bull.* 2005;28(May (5)):854–859https://www.ncbi.nlm.nih.gov/pubmed/15863893.

24. Kang SA, Park HJ, Kim MJ, Lee SY, Han SW, Leem KH. Citri Reticulatae Viride Pericarpium extract induced apoptosis in SNU-C4, human colon cancer cells. *J Ethnopharmacol.* 2005;97(February (2)):231–235https://www.ncbi.nlm.nih.gov/pubmed/15707758.

25. Lin J, Wei L, Xu W, Hong Z, Liu X, Peng J. Effect of Hedyotis Diffusa Willd extract on tumor angiogenesis. *Mol Med Rep.* 2011;4(November-December (6)):1283–1288https://www.ncbi.nlm.nih.gov/pubmed/21887465.

26. Lin J, Li Q, Chen H, Lin H, Lai Z, Peng J. Hedyotis diffusa Willd. extract suppresses proliferation and induces apoptosis via IL-6-inducible STAT3 pathway inactivation in human colorectal cancer cells. *Oncol Lett.* 2015;9(April (4)):1962–1970https://www.ncbi.nlm.nih.gov/pubmed/25789077.

27. Lu M, Wang C, Wang J. Tanshinone I induces human colorectal cancer cell apoptosis: the potential roles of Aurora A-p53 and survivin-mediated signaling pathways. *Int J Oncol.* 2016;49(August (2)):603–610https://www.ncbi.nlm.nih.gov/pubmed/27279458.

28. Su CC, Chen GW, Lin JG. Growth inhibition and apoptosis induction by tanshinone I in human colon cancer Colo 205 cells. *Int J Mol Med.* 2008;22(November (5)):613–618https://www.ncbi.nlm.nih.gov/pubmed/18949381.

29. Shih CK, Chiang W, Kuo ML. Effects of adlay on azoxymethane-induced colon carcinogenesis in rats. *Food Chem Toxicol.* 2004;42(August (8)):1339–1347https://www.ncbi.nlm.nih.gov/pubmed/15207385.

30. Li SC, Chen CM, Lin SH, Chiang W, Shih CK. Effects of adlay bran and its ethanolic extract and residue on preneoplastic lesions of the colon in rats. *J Sci Food Agric.* 2011;91(February (3)):547–552https://www.ncbi.nlm.nih.gov/pubmed/21218491.

31. Chen ST, Lee TY, Tsai TH, et al. The Traditional Chinese Medicine DangguiBuxue Tang Sensitizes Colorectal Cancer Cells to Chemoradiotherapy. *Molecules.* 2016;6(December (12)):21https://www.ncbi.nlm.nih.gov/pubmed/27929437.

32. Zhu B, Zhang QL, Hua JW, Cheng WL, Qin LP. The traditional uses, phytochemistry, and pharmacology of Atractylodes macrocephala Koidz.: a review. *J Ethnopharmacol.* 2018;226(August (18)):143–167https://www.ncbi.nlm.nih.gov/pubmed/30130541.

33. Liu H, Zhu Y, Zhang T, et al. Anti-tumor effects of atractylenolide I isolated from Atractylodes macrocephala in human lung carcinoma cell lines. *Molecules.* 2013;18(October (11)):13357–13368https://www.ncbi.nlm.nih.gov/pubmed/24172243.

34. Fu J, Ke X, Tan S, et al. The natural compound codonolactone attenuates TGF-beta1-mediated epithelial-to-mesenchymal transition and motility of breast cancer cells. *Oncol Rep.* 2016;35(January (1)):117–126https://www.ncbi.nlm.nih.gov/pubmed/26549400.

35. Sun Z, Su YH, Yue XQ. [Professor Ling Changquan's experience in treating primary liver cancer: an analysis of herbal medication]. *Zhong Xi Yi Jie He Xue Bao.* 2008;6(December (12)):1221–1225https://www.ncbi.nlm.nih.gov/pubmed/19063833.

36. Wang T, Long F, Zhang X, Yang Y, Jiang X, Wang L. Chemopreventive effects of atractylenolide II on mammary tumorigenesis via activating Nrf2-ARE pathway. *Oncotarget.* 2017;8(September (44)):77500–77514https://www.ncbi.nlm.nih.gov/pubmed/29100404.

37. Tian S, Yu H. Atractylenolide II Inhibits Proliferation, Motility and Induces Apoptosis in Human Gastric Carcinoma Cell Lines HGC-27 and AGS. *Molecules.* 2017;22(November (11))https://www.ncbi.nlm.nih.gov/pubmed/29099789.

38. Long F, Wang T, Jia P, et al. Anti-tumor effects of atractylenolide-i on human ovarian cancer cells. *Med Sci Monit.* 2017;31(January (23)):571–579https://www.ncbi.nlm.nih.gov/pubmed/28141785.

39. Wen H, Wu Z, Hu H, et al. The anti-tumor effect of pachymic acid on osteosarcoma cells by inducing PTEN and Caspase 3/7-dependent apoptosis. *J Nat Med.*

2018;72(January (1)):57–63https://www.ncbi.nlm.nih.gov/pubmed/28856634.

40. Ling H, Zhou L, Jia X, Gapter LA, Agarwal R, Ng KY. Polyporenic acid C induces caspase-8-mediated apoptosis in human lung cancer A549 cells. *Mol Carcinog.* 2009;48(June (6)):498–507https://www.ncbi.nlm.nih.gov/pubmed/18973184.

41. Cheng S, Swanson K, Eliaz I, McClintick JN, Sandusky GE, Sliva D. Pachymic acid inhibits growth and induces apoptosis of pancreatic cancer in vitro and in vivo by targeting ER stress. *PLoS One.* 2015;10(4):e0122270https://www.ncbi.nlm.nih.gov/pubmed/25915041.

42. Tin MM, Cho CH, Chan K, James AE, Ko JK. Astragalus saponins induce growth inhibition and apoptosis in human colon cancer cells and tumor xenograft. *Carcinogenesis.* 2007;28(June (6)):1347–1355https://www.ncbi.nlm.nih.gov/pubmed/17148504.

43. Tseng A, Yang CH, Chen CH, et al. An in vivo molecular response analysis of colorectal cancer treated with Astragalus membranaceus extract. *Oncol Rep.* 2016;35(February (2)):659–668https://www.ncbi.nlm.nih.gov/pubmed/26719057.

44. Lin J, Dong HF, Oppenheim JJ, Howard OM. Effects of astragali radix on the growth of different cancer cell lines. *World J Gastroenterol.* 2003;9(April (4)):670–673https://www.ncbi.nlm.nih.gov/pubmed/12679907.

45. Sun YX. Immunological adjuvant effect of a water-soluble polysaccharide, CPP, from the roots of Codonopsis pilosula on the immune responses to ovalbumin in mice. *Chem Biodivers.* 2009;6(June (6)):890–896https://www.ncbi.nlm.nih.gov/pubmed/19551730.

46. Liu X, Xu F, Wang G, Diao X, Li Y. Kanglaite injection plus chemotherapy versus chemotherapy alone for non-small cell lung cancer patients: A systematic review and meta-analysis. *Curr Ther Res Clin Exp.* 2008;69(October (5)):381–411https://www.ncbi.nlm.nih.gov/pubmed/24692815.

47. Guo Q, Li J, Lin H. Effect and molecular mechanisms of traditional chinese medicine on regulating tumor immunosuppressive microenvironment. *Biomed Res Int.* 2015;2015:261620https://www.ncbi.nlm.nih.gov/pubmed/26161392.

48. Qin K, Zheng L, Cai H, et al. Characterization of Chemical Composition of Pericarpium Citri Reticulatae Volatile Oil by Comprehensive Two-Dimensional Gas Chromatography with High-Resolution Time-of-Flight Mass Spectrometry. *Evid Based Complement Alternat Med.* 2013;2013:237541https://www.ncbi.nlm.nih.gov/pubmed/23710215.

49. Chen Z, Li S, Wang X, Zhang CL. Protective effects of Radix Pseudostellariae polysaccharides against exercise-induced oxidative stress in male rats. *Exp Ther Med.* 2013;5(April (4)):1089–1092https://www.ncbi.nlm.nih.gov/pubmed/23596474.

50. Lee HZ, Bau DT, Kuo CL, Tsai RY, Chen YC, Chang YH. Clarification of the phenotypic characteristics and anti-tumor activity of Hedyotis diffusa. *Am J Chin Med.* 2011;39(1):201–213https://www.ncbi.nlm.nih.gov/pubmed/21213409.

51. Li YL, Zhang J, Min D, Hongyan Z, Lin N, Li QS. Anticancer effects of 1,3-Dihydroxy-2-Methylanthraquinone and the ethyl acetate fraction of Hedyotis Diffusa willd against HepG2 carcinoma cells mediated via apoptosis. *PLoS One.* 2016;11(4):e0151502https://www.ncbi.nlm.nih.gov/pubmed/27064569.

52. Liu Z, Liu M, Liu M, Li J. Methylanthraquinone from Hedyotis diffusa WILLD induces Ca(2+)-mediated apoptosis in human breast cancer cells. *Toxicol In Vitro.* 2010;24(February (1)):142–147https://www.ncbi.nlm.nih.gov/pubmed/19686834.

53. Lin M, Lin J, Wei L, et al. Hedyotis diffusa Willd extract inhibits HT-29 cell proliferation via cell cycle arrest. *Exp Ther Med.* 2012;4(August (2)):307–310https://www.ncbi.nlm.nih.gov/pubmed/23139718.

54. Lin J, Chen Y, Wei L, et al. Hedyotis Diffusa Willd extract induces apoptosis via activation of the mitochondrion-dependent pathway in human colon carcinoma cells. *Int J Oncol.* 2010;37(November (5)):1331–1338https://www.ncbi.nlm.nih.gov/pubmed/20878081.

55. Zhang X, Cai Y, Wang L, Liu H, Wang X. Optimization of processing technology of Rhizoma Pinelliae Praeparatum and its anti-tumor effect. *Afr Health Sci.* 2015;15(March (1)):101–106https://www.ncbi.nlm.nih.gov/pubmed/25834537.

56. Zu G, Wang H, Wang J, Dou Y, Zhao W, Sun Y. Rhizoma Pinelliae trypsin inhibitor separation, purification and inhibitory activity on the proliferation of BGC-823 gastric adenocarcinoma cells. *Exp Ther Med.* 2014;8(July (1)):248–254https://www.ncbi.nlm.nih.gov/pubmed/24944630.

57. Yoon JS, Seo JC, Han SW. Pinelliae Rhizoma herbal-acupuncture solution induced apoptosis in human cervical cancer cells, SNU-17. *Am J Chin Med.* 2006;34(3):401–408https://www.ncbi.nlm.nih.gov/pubmed/16710889.

58. Juan LK-l.CB-s.HXL. *Anti-tumor activities of extracts from the medicinal plants Pinellia ternata and Pinellia pedatisecta.* 2009 3rd International Conference on Bioinformatics and Biomedical Engineering. 2009; 2009.

59. Zhao Y, Feng LM, Liu LJ, Zhang X, Zhao RZ. Clerosterol from vinegar-baked radix bupleuri modifies drug transport. *Oncotarget.* 2017;8(March (13)):21351–21361https://www.ncbi.nlm.nih.gov/pubmed/28423482.

60. Chow LW, Loo WT, Sham JS, Cheung MN. Radix bupleuri containing compound (KY88 liver-livo) induces apoptosis and production of interleukin-4 and tumor necrosis factor-alpha in liver cancer cells in vitro. *Am J Chin Med.*

61. 2004;32(2):185–193https://www.ncbi.nlm.nih.gov/pubmed/15315257.

61. Kok LD, Wong CK, Leung KN, Tsang SF, Fung KP, Choy YM. Activation of the anti-tumor effector cells by Radix bupleuri. *Immunopharmacology.* 1995;30(June (1)):79–87https://www.ncbi.nlm.nih.gov/pubmed/7591716.

62. Hsu YL, Kuo PL, Weng TC, Yen MH, Chiang LC, Lin CC. The antiproliferative activity of saponin-enriched fraction from Bupleurum Kaoi is through Fas-dependent apoptotic pathway in human non-small cell lung cancer A549 cells. *Biol Pharm Bull.* 2004;27(July (7)):1112–1115https://www.ncbi.nlm.nih.gov/pubmed/15256750.

63. Wang BF, Dai ZJ, Wang XJ, et al. Saikosaponin-d increases the radiosensitivity of smmc-7721 hepatocellular carcinoma cells by adjusting the g0/g1 and g2/m checkpoints of the cell cycle. *BMC Complement Altern Med.* 2013;12(October (13)):263https://www.ncbi.nlm.nih.gov/pubmed/24119370.

64. Wang Z, Yu G, Liu Z, et al. Paeoniflorin inhibits glioblastoma growth in vivo and in vitro: a role for the Triad3A-dependent ubiquitin proteasome pathway in TLR4 degradation. *Cancer Manag Res.* 2018;10:887–897https://www.ncbi.nlm.nih.gov/pubmed/29740218.

65. Fang S, Zhu W, Zhang Y, Shu Y, Liu P. Paeoniflorin modulates multidrug resistance of a human gastric cancer cell line via the inhibition of NF-kappaB activation. *Mol Med Rep.* 2012;5(February (2)):351–356https://www.ncbi.nlm.nih.gov/pubmed/22051979.

66. Li P, Zhang ZM, Li T, et al. Monoterpene derivatives from the roots of Paeonia lactiflora and their anti-proliferative activity. *Fitoterapia.* 2014;98(October):124–129https://www.ncbi.nlm.nih.gov/pubmed/25068201.

67. Washida K, Itoh Y, Iwashita T, Nomoto K. Androgen modulators from the roots of Paeonia lactiflora (paeoniae radix) grown and processed in Nara prefecture, Japan. *Chem Pharm Bull (Tokyo).* 2009;57(September (9)):971–974https://www.ncbi.nlm.nih.gov/pubmed/19721258.

68. Lee SM, Li ML, Tse YC, et al. Paeoniae Radix, a Chinese herbal extract, inhibit hepatoma cells growth by inducing apoptosis in a p53 independent pathway. *Life Sci.* 2002;71(September (19)):2267–2277https://www.ncbi.nlm.nih.gov/pubmed/12215374.

69. Huang F, Li S, Lu X, Liu A, Du G, Shi G. Two glutathione S-transferase inhibitors from Radix Angelicae sinensis. *Phytother Res.* 2011;25(February (2)):284–289https://www.ncbi.nlm.nih.gov/pubmed/20665471.

70. Chen YL, Jian MH, Lin CC, et al. The induction of orphan nuclear receptor Nur77 expression by n-butylenephthalide as pharmaceuticals on hepatocellular carcinoma cell therapy. *Mol Pharmacol.* 2008;74(October (4)):1046–1058https://www.ncbi.nlm.nih.gov/pubmed/18577687.

71. Chiu SC, Chen SP, Huang SY, et al. Induction of apoptosis coupled to endoplasmic reticulum stress in human prostate cancer cells by n-butylidenephthalide. *PLoS One.* 2012;7(3):e33742https://www.ncbi.nlm.nih.gov/pubmed/22470469.

72. Lin PC, Chen YL, Chiu SC, et al. Orphan nuclear receptor, Nurr-77 was a possible target gene of butylidenephthalide chemotherapy on glioblastoma multiform brain tumor. *J Neurochem.* 2008;106(August (3)):1017–1026https://www.ncbi.nlm.nih.gov/pubmed/18419761.

73. Wei CW, Lin CC, Yu YL, et al. n-Butylidenephthalide induced apoptosis in the A549 human lung adenocarcinoma cell line by coupled down-regulation of AP-2alpha and telomerase activity. *Acta Pharmacol Sin.* 2009;30(September(9)):1297–1306https://www.ncbi.nlm.nih.gov/pubmed/19701232.

74. Chiu SC, Chen TL, Huang SY, et al. Potential therapeutic effects of N-butylide-nephthalide from Radix Angelica Sinensis (Danggui) in human bladder cancer cells. *BMC Complement Altern Med.* 2017;17(December (1)):523https://www.ncbi.nlm.nih.gov/pubmed/29207978.

75. Gao H, Sun W, Zhao W, et al. Total tanshinones-induced apoptosis and autophagy via reactive oxygen species in lung cancer 95D Cells. *Am J Chin Med.* 2015;43(6):1265–1279https://www.ncbi.nlm.nih.gov/pubmed/26394653.

76. Zhang J, Wang J, Jiang JY, Liu SD, Fu K, Liu HY. Tanshinone IIA induces cytochrome c-mediated caspase cascade apoptosis in A549 human lung cancer cells via the JNK pathway. *Int J Oncol.* 2014;45(August (2)):683–690https://www.ncbi.nlm.nih.gov/pubmed/24888720.

77. Wu CY, Cherng JY, Yang YH, et al. Danshen improves survival of patients with advanced lung cancer and targeting the relationship between macrophages and lung cancer cells. *Oncotarget.* 2017;8(October (53)):90925–90947https://www.ncbi.nlm.nih.gov/pubmed/29207614.

78. Lin YY, Lee IY, Huang WS, et al. Danshen improves survival of patients with colon cancer and dihydroisotanshinone I inhibit the proliferation of colon cancer cells via apoptosis and skp2 signaling pathway. *J Ethnopharmacol.* 2017;14(September (209)):305–316.

79. Guo Y, Nie Q, MacLean AL, Li Y, Lei J, Li S. Multiscale Modeling of Inflammation-Induced Tumorigenesis Reveals Competing Oncogenic and Oncoprotective Roles for Inflammation. *Cancer Res.* 2017;77(November (22)):6429–6441https://www.ncbi.nlm.nih.gov/pubmed/28951462.