Journal of Traditional and Complementary Medicine 15 (2025) 215–228

Contents lists available at ScienceDirect

Journal of Traditional and Complementary Medicine

journal homepage: www.elsevier.com/locate/jtcme

A review of recent artificial intelligence for traditional medicine

Chengbin Hou [a,1], Yanzhuo Gao [b,1], Xinyu Lin [a], Jinchao Wu [a], Ning Li [a], Hairong Lv [c], William Cheng-Chung Chu [a,*]

[a] School of Computing and Artificial Intelligence, Fuyao University of Science and Technology, Fujian, 350109, China
[b] Maynooth International Engineering College, Fuzhou University, Fuzhou, 350108, China
[c] MOE Key Laboratory of Bioinformatics, BNRIST Bioinformatics Division, Department of Automation, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Keywords:
Artificial intelligence
Traditional medicine
AI for medicine
AI for health
Machine learning
Deep learning
Large language models

ABSTRACT

Traditional Medicine (TM) has played a crucial role in global healthcare due to its long history and holistic approach. Artificial Intelligence (AI) has emerged as a revolutionary technology, offering exceptional capabilities in areas such as data mining, pattern recognition, and decision-making. The integration of Artificial Intelligence for Traditional Medicine (AITM) presents a promising frontier in advancing medicine and healthcare. In this review, we explore AITM from two perspectives: recent AI techniques and TM applications. Specifically, we investigate how Machine Learning, Deep Learning, and Large Language Models are applied to TM, covering applications such as diagnosis (before, during, after) and research (drug research, structured knowledge, data analysis). By leveraging advanced algorithms and models, AI can improve decision-making efficiency, optimize diagnosis accuracy, enhance patient experience, and reduce costs. We anticipate this review can bridge the gap between AI and TM communities. And the goal is to foster collaboration and innovation between both communities, enabling them to exploit the state-of-the-art AI techniques to advance TM diagnosis and research, ultimately contributing to the enhancement of human health.

List of Abbreviations:

TM	Traditional Medicine
TCM	Traditional Chinese Medicine
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
LLM or LLMs	Large Language Models
SVM	Support Vector Machine
DT	Decision Tree
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
BERT	Bidirectional Encoder Representations from Transformer
GPT	Generative Pre-trained Transformer
ANN	Artificial Neural Network
GNN	Graph Neural Network
LR	Logistic Regression
KNN	K-Nearest Neighbors
RF	Random Forest
NB	Naive Bayesian
NER	Named Entity Recognition

1. Introduction

Traditional Medicine (TM), spanning centuries across various cultures and civilizations, comprises a wide range of healing practices that have been passed down through generations. According to the definition by the World Health Organization, TM is “the sum total of the knowledge, skill, and practices based on the theories, beliefs, and experiences indigenous to different cultures, whether explicable or not, used in the maintenance of health as well as in the prevention, diagnosis, improvement or treatment of physical and mental illness”.[1] Distinguished from other medicines, TM often focuses on holistic approaches, addressing not only the physical symptoms but also the emotional, spiritual, and environmental factors that contribute to health and illness. Nowadays, TM still remains crucial, especially in developing countries and rural areas, due to its accessibility, affordability, and alignment with local cultural values.[2,3] The importance of TM lies in its historical and cultural values, and more importantly in its potential to offer the alternative solution to other medical solutions.[3–5]

Artificial Intelligence (AI) can be considered as a branch of computer

* Corresponding author.
E-mail addresses: chengbin.hou10@foxmail.com (C. Hou), lvhairong@tsinghua.edu.cn (H. Lv), chu@fyust.org.cn (W.C.-C. Chu).
[1] Co-first authors.

https://doi.org/10.1016/j.jtcme.2025.02.009
Received 15 November 2024; Received in revised form 30 December 2024; Accepted 20 February 2025
Available online 21 February 2025

science that simulates human intelligence such as learning, reasoning, and decision-making. Its most significant and distinctive capabilities lie in analyzing complex data, learning from experience, adapting to new inputs, and making data-driven decisions.[6] The rapid progress of AI has fueled innovation in a multitude of areas, including but not limited to autonomous vehicles, manufacturing, education, scientific research, and healthcare.

Regarding Artificial Intelligence for Traditional Medicine (AITM), AI techniques hold the potential to deeply participate in and even reshape several aspects of TM.[7–16] For instance, support vector machines and decision trees have been exploited for quantitative analysis of diseases or treatment data; convolutional neural networks have been widely applied to classify the categories of images in herbal medicine or tongue diagnosis; large language models have been employed to handle medical texts for question answering or identifying key words. By leveraging advanced algorithms and models, AI can assist TM in improving decision-making efficiency, optimizing diagnosis accuracy, enhancing patient experience, and reducing costs.

In this review, we intend to cover AITM works from two perspectives: the recent AI techniques and TM applications. Specifically, recent AI techniques mainly include Machine Learning (ML), Deep Learning (DL), and Large Language Models (LLM or LLMs), while TM applications contain diagnosis (before, during, after) and research (drug research, structured knowledge, data analysis). To start with, the terminology of the recent techniques, belonging to ML, DL, and LLM, is combined respectively with TM as the key words for search in Google Scholar. A few examples of key words for "ML + TM" are "machine leaning + traditional medicine", "support vector machine + traditional medicine", and "decision tree + traditional medicine". After that, all these works are skimmed and the key information is recorded. We then categorize them based on the keynotes and the aforementioned two perspectives.

There have been many valuable review papers concerning AI and medicine,[7–49] while these papers relatively less focus on AITM.[7–16] The existing review papers mostly cover the works using ML and DL techniques for TM diagnosis and research. However, the recent breakthrough AI technique of LLM[50,51] has not yet been covered in the context of TM. To fill the gap, this review paper aims to include the recent AI techniques of ML, DL, and LLM for both diagnosis and research applications. Unlike previous review papers, we organize AITM works based on the recent AI techniques and TM applications, so that the researchers from AI community can understand where AI techniques have been applied in TM, and the researchers from medicine community are aware of what and how AI techniques can be applied in TM. The goal is to promote both communities to leverage the state-of-the-art AI techniques to advance TM diagnosis and research, ultimately contributing to the enhancement of human health.

The rest of the paper is structured as follows. Section 2 summarizes the related review papers in AI and medicine. Next, we present an overview of the AITM pipeline and the taxonomy of AITM research works in Section 3. After that, the specific research works for ML + TM, DL + TM, and LLM + TM are elaborated in Sections 4,5, and 6 respectively based on their applications. Section 7 then discusses the challenges and future directions. And the conclusion is finally drawn in the last section.

2. Related review papers

The integration of AI in the medical discipline has been transformative, offering novel solutions for both traditional and non-traditional medicine. This section attempts to provide a detailed analysis of previous review papers on the application of AI in traditional and non-traditional medicine. As shown in Table 1, these reviews are organized around two dimensions: the medical applications including diagnosis and research, and the AI techniques that drive recent advancements. With the continuous development of AI in the medical field, the number of related reviews has also been steadily increasing,

Table 1
Summary of previous review papers on the application of AI in traditional and non-traditional medicine.

Publications (Counts)	Focus on Traditional Medicine?	Applications		AI Techniques		
		Diagnosis	Research	ML	DL	LLM
[12] (1)	✓		✓	✓	✓	
[16] (1)	✓	✓		✓	✓	
[7–11,13–15] (8)	✓	✓	✓	✓	✓	
[35] (1)		✓	✓		✓	✓
[36] (1)		✓	✓	✓	✓	✓
[37,48,52–54] (5)		✓		✓	✓	
[17–34,49] (19)		✓	✓	✓	✓	
[39,40,42–45] (6)			✓	✓	✓	
[38,41,46,47] (4)						

offering a series of useful review papers.[7–49,52–54] Although the reviews specifically targeting TM are relatively sparse,[7–16] TM has demonstrated unique value in auxiliary treatment and preventive healthcare. Herbs, acupuncture, and Tai Chi in TM have been shown to effectively treat chronic diseases (such as rheumatoid arthritis and fibromyalgia), refractory diseases (such as migraines and irritable bowel syndrome), as well as major diseases like cancer and acute promyelocytic leukemia.[15] Recently, AI has made significant breakthroughs in acupuncture, particularly in acupoint selection, prescription, technique identification, and efficacy prediction. Ref. 13 highlighted how AI driven data mining methods have identified key acupoint combinations for treating various diseases, offering a scientific basis for clinical acupoint prescriptions. Additionally, the application of AI-based DL and cloud computing has attracted widespread attention in healthcare, particularly for identifying potential active ingredients, targets, and pathways of single herbs or formulas in TM, as well as optimizing disease diagnosis and treatment models. For example, Ref. 8 illustrated how AI-assisted DL and cloud computing have been applied in both Western medicine and TM for diagnosing and treating different stages of rheumatoid arthritis.

From the application viewpoint, diagnosis and research are two essential medical aspects. We define diagnosis as the process of a patient undergoes when seeking medical assistance, including the stages before, during, and after diagnosis. Medical research, on the other hand, primarily contains drug research, knowledge structuring, and data analysis. According to Table 1, most reviews cover both diagnosis and research.[7–11,13–15,17–36,49] However, there are also a small number of reviews that focus solely on either diagnosis or research.[12,16,37,39,40,42–45,48,52–54] Reviews that focus on diagnosis often discuss AI in specific clinical applications. For example, Ref. 37 focused on AI-assisted endoscopic examination, discussing its application in gastrointestinal diseases. Ref. 48 discussed the application of AI in diagnosing various diseases, highlighting its potential to improve accuracy, accelerate detection, and enhance treatment efficacy, thereby increasing the chances of patient recovery. AI has also made progress in sports medicine, especially in the diagnosis of sports injuries. For example, Ref. 53 reviewed methods for assessing the risk of sports injuries. The ESSKA AI Working Group has promoted applications in orthopaedics, Ref. 52 reviewing the use of ML in sports injury prediction, and Ref. 54 highlighting the application of digital twin technology in orthopaedic surgery and postoperative prediction. Reviews in research usually focus on drug research, Ref. 40,42–44 discussing various applications of AI-assisted drug design (such as virtual screening, property prediction, and clinical trials), as well as the challenges and opportunities it brings to the pharmaceutical industry. Apart from medical diagnosis and research, there are also other interesting AI-related medical topics such as privacy and interpretability,[38] historical development,[41] ethics,[46] and education.[47]

From the viewpoint of recent AI techniques, a large number of reviews have thoroughly surveyed the applications of ML and DL techniques in medicine or healthcare.[7–11,13–45] Most of them employed ML and/or DL to improve the analysis of medical images, patient data and genetic information, changing clinical decision-making.[17,19–36] For instance, Ref. 20 mentioned that radio graphic images, pathological slides, and electronic medical records are being evaluated through machine learning, thereby assisting in the diagnosis and treatment of patients and enhancing the capabilities of doctors. Ref. 22 stated that AI will enhance precision medicine decision-making, particularly in handling nongenomic and genomic determinants by integrating patient symptoms, clinical history, and lifestyle information. Ref. 33 elaborated on the contribution of AI in precision medicine and genomic medicine, where researchers can leverage ML and DL techniques along with extensive datasets to usher in a new era of effective genomic medicine. It is worth emphasizing that LLMs have emerged as the cutting-edge AI technique with significant impacts across various fields including medicine. There are two short reviews concerning LLMs and medicine. Ref. 35 addressed the current efforts of applying LLMs in medicine, but overlooked other important AI techniques such as ML and DL. Ref. 31 discussed the impact of the COVID-19 pandemic in accelerating the use of AI in areas, such as telemedicine and chatbots, to enhance accessibility and improve medical education.

To the best of our knowledge, LLMs have not yet been systematically covered in the related reviews, especially with the focus on TM. Distinguished from the previous reviews, this review not only includes the recent ML and DL techniques, but also delves into the emerging technique of LLMs. Meanwhile, this review also considers the applications of the recent AI techniques in both TM diagnosis and research.

3. Overview and preliminary

As illustrated in Fig. 1, the overview of AI and TM can be organized into four parts: data, techniques, tasks, and applications. First, the modality of TM data could be native images, native texts, and other features, which are either directly or preprocessed as the input to AI models. Second, the recent AI techniques can be generally divided into three groups, i.e., ML, DL, and LLM; each group further contains more specific AI techniques that can be adopted to handle TM data and perform AI tasks. Third, there are several typical types of general AI tasks, e.g., classification tasks, to formulate a TM problem into an AI problem, e.g., classifying the category of herb images. Fourth, with the paradigm of formulating TM problems into AI problems, AI techniques can be then applied to TM data and accordingly support various TM applications such as the traditional medical diagnosis and drug research.

3.1. TM data

Data is crucial in medical diagnosis and research, as both human experts and AI models rely on data for learning and decision-making. Unlike human experts, AI models often require much more data for training the model, which can then be applied to make predictions based on the patterns learned during the model training phase. Considering current AITM works, the modalities of TM data mainly include but not limited to feature-based data, image data, and textual data.

For the less recent AI techniques, these data are often preprocessed into the so-called embeddings such that each sample becomes a data point in a high dimensional vector space. For example, we can adopt one-hot encoding[55,56] to preprocess the original features or the textural data; we may flatten out the image pixels as one-dimensional vector. For the more recent AI techniques, these data can be directly fed into the models without too much feature engineering. For instances, some recent AI techniques can effectively extract key features from the raw images in an end-to-end supervised learning paradigm[57,58]; some very recent AI techniques could even learn from the raw data without human labeling in an end-to-end unsupervised or self-supervised learning paradigm.[50,51]

3.2. AI techniques

The recent AI techniques mainly cover three groups: ML, DL, and LLM. As depicted in the Venn diagram in Fig. 1, LLM is a subset of DL; DL is the subset of ML; all the three groups are the subset of AI, while AI would contain other techniques not in the three groups. For the three groups covered in this work, we first consider the early learning model, e.g., Support Vector Machine (SVM)[59] and Decision Tree (DT),[60] into the ML group. Second, the more recent learning model allowing deeper layers connected together, e.g., Convolutional Neural Networks (CNN)[57] and Recurrent Neural Networks (RNN),[61] is regarded as the DL group. And finally, the learning model typically with over millions of parameters, e.g., Bidirectional Encoder Representations from Transformers (BERT)[50] and Generative Pre-trained Transformer (GPT),[51] is treated as the LLM group. It might be worth mentioning that the years of 2015 and 2019 are based on the two landmark papers in which Ref. 62 presented the review of deep learning (over 85k citations) and Ref. 50 developed the BERT model (over 115k citations), although the related techniques of DL and LLM might be proposed earlier than the years of 2015 and 2019.

Machine Learning (ML) refers to statistical models that can learn

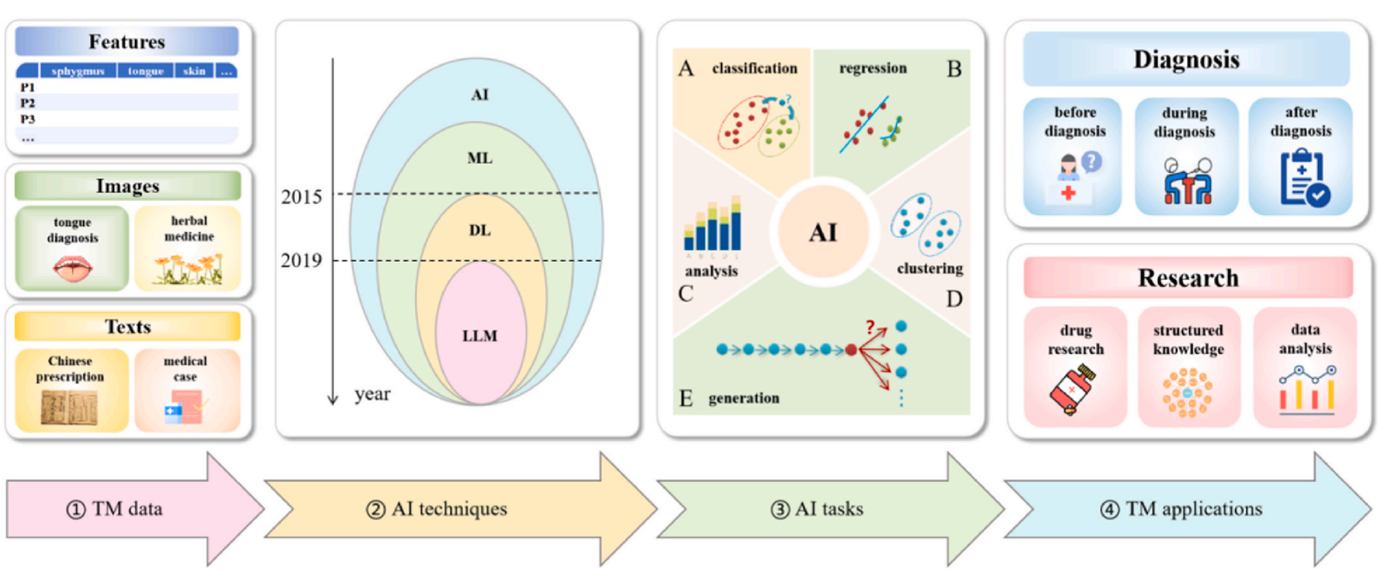

Fig. 1. Overview of Artificial Intelligence (AI) and Traditional Medicine (TM): data, techniques, tasks, and applications.

underlying patterns from data, and generalize to unseen data without following explicit instructions. And in this work, ML is used to represent a set of early ML techniques including but not limiting to Logistic Regression (LR), Support Vector Machine (SVM), Decision Tree (DT), and shallow Artificial Neural Network (ANN). For more ML techniques, readers are encouraged to consult the early ML review[63] as well as the ready to use Python library https://scikit-learn.org/. Mathematically, we describe the early ML models as follows:

$$y = f(\phi(\mathbf{x})) \tag{1}$$

where $\mathbf{x}$ is the original input features; ϕ (.) could be a mapping for preprocessing the input features; f (.) is the function that maps its input to the output y. It is worth mentioning that f could be either parametric (e.g., LR and ANN) or non-parametric (e.g., SVM and DT), and the parametric f contains a fixed set of trainable parameters.

Deep Learning (DL) is a subset of ML that typically employs the deep neural networks with many layers or even more than hundreds of layers. The layers or modules to establish deep learning models include fully connected layer, Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Graph Neural Network (GNN), attention layer, and so on. Note that, CNN is preferred to deal with visual data like images, while RNN is good at handling sequential data like texts. For more DL techniques, we refer readers to the recent DL review[62] as well as the ready to use Python libraries such as https://pytorch.org/ and https://www.tensorflow.org/. To be more specific, the formulation of DL models can be expressed as:

$$y = f^l\left(\cdots f^1\left(f^0(\mathbf{x};\theta_0);\theta_1\right);\cdots\theta_l\right) \tag{2}$$

where $\mathbf{x}$ is the original input features; θ denotes the model parameters and is used to parameterize the corresponding function; f is the parameterized function typically with non-linear activation, and f could be CNN, RNN, or other type of layers; the superscript $0, 1, \ldots l$ indicates the number of layers. As the layer goes deeper, the more high-level features would be extracted from the initial $\mathbf{x}$, but meanwhile the DL model would become harder to train. Unlike traditional ML formulated in Equation (1) that the original input $\mathbf{x}$ often requires human experts for feature engineering via $\phi(\mathbf{x})$, DL empowers deep networks to automatically extract useful features in the end-to-end learning fashion.

Large Language Model (LLM) is the deep learning model typically containing over millions of parameters, and the most popular building block of LLM is transformer.[64] LLM is trained on the massive textual data in a self-supervised fashion. Essentially, LLM is a statistical probabilistic model that assigns probabilities to sequences of words. It aims to maximize the co-occurrence probabilities given the sequences of words produced by humans, so that it can learn from human natural language. The probability of a sequence of words can be calculated by:

$$P(w_1, w_2, \cdots, w_T) = \prod_{t=1}^{T} P(w_t \mid w_1, w_2, \cdots, w_{t-1}) \tag{3}$$

where $P(w_t \mid w_1, w_2, \cdots, w_{t-1})$ is the conditional probability of the word w_t given its previous context $w_1, w_2, \cdots, w_{t-1}$; the joint probability of a sequence of words $P(w_1, w_2, \cdots, w_T)$ can be obtained by the product of the conditional probabilities over the sequence of words. We can estimate the probabilities using parametric neural networks $f(w_1, w_2, \cdots, w_T; \theta)$[65]. Thanks to the development of the attention mechanism and transformer architecture,[64] LLM can goes deeper with even millions or billions parameters θ and thereby enhancing the ability to model and understand human natural language. However, as training a such huge model from scratch is intractable for most people due to unaffordable computational resources, it is recommended to directly use the trained LLM or sometime fine-tune a trained LLM on, e.g., TM textural data for TM purpose. The well trained LLMs can be found at https://huggingface.co/, and there have been already successful examples,[2],[3],[4] of LLMs specifically tailored for TM.

3.3. AI tasks

To properly formulate TM problems into AI problems, it is suggested to understand some common AI tasks. One of the standard tasks is classification,[66] which aims to classify a given sample into the predefined classes based on its features. One of widely-used loss functions in classification is cross-entropy loss, which compares the predicted label given by the model and the ground-truth label annotated by humans, and accordingly tries to minimize the errors over all samples. Another standard task is regression,[67] which attempts to predict the continuous score for a given sample based on its features. One basic loss function in regression is mean-squared error loss, and the objective for training the model is to minimize the errors overall samples between the predicted scores and ground-truth scores.

Apart from classification and regression, the generation task has recently become popular, which intends to generate natural language imitating humans for question answering and even reasoning. The objective is to maximize the occurrence probability of the masked word given its surrounding words[50] or the next word given its previous words,[51] so as to train the neural probabilistic language model as shown in Equation (3) with a huge number of parameters over massive texts. By doing so, the underlying text patters and human knowledge behind texts would be compressed into the model parameters of LLM.

There are also other AI tasks such as clustering and factor analysis. Note that, these common AI tasks can be used to transform TM problems into AI problems, which facilitates utilizing AI techniques to solve TM problems, and thereby supports various TM applications.

3.4. TM applications

In general, TM applications can be divided into two groups. The first one is related to TM diagnosis, which covers the full cycle of patients seeking medical help from clinicians or doctors, including before diagnosis, during diagnosis, and after diagnosis. Specifically in this work, the "before diagnosis" refers to the point before the first diagnosis meeting between patients and doctors. The "after diagnosis" is the point after patients receiving his/her doctor's advice in which patients are not required to make diagnosis again in a short period of time.

Another one is about TM research, which does not directly deal with the specific illness of a patient, consisting of drug research, structured knowledge, and data analysis. It is noted that collecting and building structured knowledge or structured data would serve as the basis to support many other TM applications. For examples, the structured knowledge can assist LLM to better respond diagnostic queries; having more structured data available can enhance model training and thus lead to better performance.

3.5. Taxonomy of AITM works

Overall, the AITM works are organized according to two perspectives: recent AI techniques (ML, DL, and LLM) and TM applications (diagnosis: before, during, after; research: drug, structured knowledge, data analysis). The flowchart of how to organize these works is illustrated in Fig. 2. And we follow this workflow in Sections 4,5, and 6.

[2] https://huggingface.co/TCMLLM/Lingdan-13B-PR.
[3] https://huggingface.co/CMLM/ZhongJing-2-1_8b.
[4] https://huggingface.co/tyang816/MedChatZH.

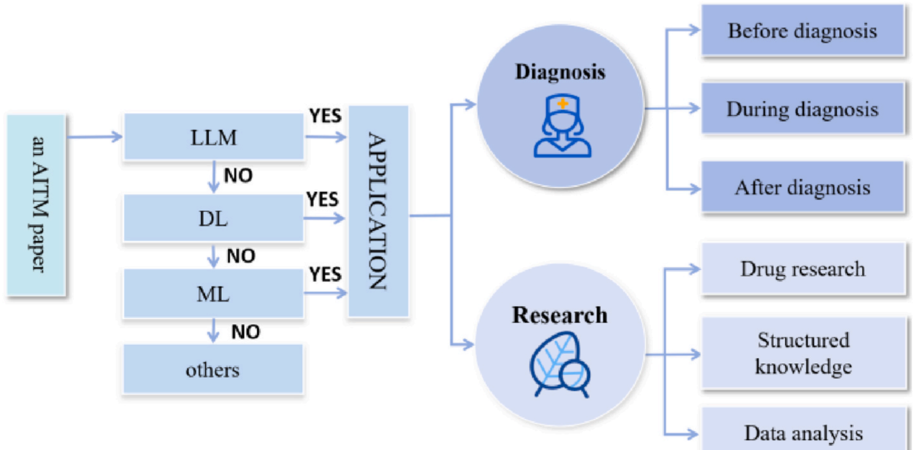

Fig. 2. Flowchart for taxonomy.

Table 2
Overview and statistics of ML Applications in Traditional Medicine.

Application	Description	Model	Data (Size)	Publications (Year)
Before diagnosis	Identifying TCM Constitution Types and Analyzing Their Link to Obesity	LR	CCMQ Responses (18805)	[68] (2010)
	Classification of Body Constitution Types from the BCQ	DT	BCQ Responses (108)	[69] (2022)
During diagnosis	TCM Model for Syndrome Differentiation Formulas of Chronic Hepatitis B	DT	Patient Records (555)	[70] (2011)
	Classification of Lip Images for TCM Lip Diagnosis	SVM	Lip Images (257)	[71] (2012)
	Patient Classification of Hypertension	SVM, KNN	Hypertension Clinical Cases (904)	[72] (2015)
	Analyzing Clinical Records for Diagnosis and Acupoint Recommendations	Ensemble Model	Clinical Records (2,835)	[73] (2015)
	Firefly and Simulated Annealing Algorithms for Clustering Prescriptions	K-means	CM Prescriptions (1,100)	[74] (2016)
	Predicting Prescriptions for Disease Treatment	SVM	Abstract Text (48,243)	[75] (2017)
	Warfarin Dose Prediction Based on Patient Clinical Variables	Ensemble Model	Patient Clinical Data (289)	[76] (2018)
	Classifying Chinese Medicine Prescriptions Using HIN Modeling	K-means, etc.	ChP and TCM clinical cases (3K+)	[77] (2019)
	Classification of Pulse Signals	SVM, etc.	Pulse Data (1,812)	[78] (2021)
	TCM Constitution Discrimination	RF, DT	Examination Results (441)	[79] (2022)
	Multi-Label Classification for TCM Prescription Generation in Lung Cancer	SVM	TCM Prescriptions (10,000)	[80] (2022)
Drug research	Identification of Drug-Like Compounds Through Drug-Likeness Prediction	NB	MDDR, ACD, TCMCD	[81] (2012)
	Discrimination of Herbal Medicines Using Terahertz Spectroscopy	SVM, DT, RF	Traditional Herbal Medicine (1200)	[82] (2017)
	Integrating Multiple Classic Classifiers for Ginseng Identification	Ensemble Model	Ginseng Samples (315)	[83] (2017)
	Classification of Mineral TCMs Based on Spectral Data	SVM	Spectral Data (108)	[84] (2019)
	Identification of Neuroprotective Compounds in XXMD	KNN, RF, etc.	Herbal Compounds in XXMD (1,484)	[85] (2019)
	Clustering Herbal Medicines by Automated Cluster Center Selection	K-means, etc.	Prescriptions (10,000+)	[86] (2019)
	Classification of Chinese Herbs Efficacy Based on Node Similarity	K-means	Prescriptions (22,570) and Herbs (1,956)	[87] (2020)
	Screening and Discovery of Antibacterial Herbal Medicines	ANN, RF, etc.	Antibacterial TCMs (904)	[88] (2021)
	Discrimination of Herbal Medicines Based on FT-IR Spectra	DT, RF, ANN	Mutong Samples (164)	[89] (2021)
	Daily Intake Estimation of Hepatotoxic Compounds	SVM	Hepatotoxicity of 17 NP-TCMs	[90] (2021)
	Screening Hepatotoxic Compounds in TCM and Western Medicine	SVM, RF, NB, etc.	Hepatotoxic Compounds (2,035)	[91] (2022)
	Screening Western and TCM Combinations for RA Treatment	KNN, SVM, RF, etc.	DrugCombDB, DrugBank, TCMSP etc.	[92] (2023)
Data analysis	Extracting Syndrome Differentiation Rules for Post-Hepatic Cirrhosis	DT	Patient Records (293)	[93] (2008)
	Determining Risk Factors of Osteoporosis	RF, SVM	Questionnaires (998)	[94] (2012)
	Prediction of Cardiotoxicity in TCM Compatibility	SVM	Herb Groups	[95] (2016)
	Identification of Articles on TM	SVM	MEDLINE (189,674)	[96] (2017)
	Quality Classification of KuDieZi Injection	SVM	KuDieZi Injection (50 batches)	[97] (2017)
	Quantitative Prediction of Glycyrrhizic Acid Content in Chinese Herbal Medicine Glycyrrhiza Glabra Based on Four Extraction Parameters	GA-PBNN, etc.	Glycyrrhizic Acid Extraction Data (30)	[98] (2018)
	Prediction of Incompatible Herb Pairs in TCM	LR, SVM, etc.	IHP datasets (528)	[99] (2019)
	Predict Meridian of Herbs from Molecular Features	SVM, DT, RF, KNN	TCMID (18,140)	[100] (2019)
	Exploring Key Aspects of TCM Diagnosis and Treatment for Insomnia	RF	Outpatient Cases (654)	[101] (2021)
	Exploring the Most Frequently Used Drugs for Treating Acute Pancreatitis	RF	Compounds for AP Treatment (516)	[102] (2021)

4. ML on traditional medicine

As an early AI technique, ML can develop task-specific models tailored for various applications in TM. Through our investigation, we found that classification tasks of ML are the most widely applied in TM, while the applications of regression tasks are relatively rare. This is primarily because TM often relies on qualitative judgments to study problems rather than quantitative analyses. Additionally, the clustering tasks regarding ML are commonly used to discover or identify samples with similar characteristics. Furthermore, many studies utilize ML techniques for data analysis to enhance the interpretability of TM in practical applications. In this section, we focus on the specific applications of ML in TM, as shown in Table 2, which can be divided into two main parts: diagnosis and research.

4.1. ML applications in diagnosis

The applications of AI in diagnosis consist of three stages: before diagnosis, during diagnosis, and after diagnosis. Regarding the before diagnosis stage, ML mainly classifies patients based on questionnaires, aiding subsequent diagnosis and treatment. For instance, Ref. 68 and Ref. 69 employed the standardized Constitution in Chinese Medicine Questionnaire (CCMQ) and Body Constitution Questionnaire (BCQ), respectively. Both questionnaires contain multiple questions with multi-level scoring to determine Traditional Chinese Medicine (TCM) constitution types (e.g., Yang-Xu, Yin-Xu, and Stasis types). These data were then used to train Logistic Regression (LR) and Decision Tree (DT) models for the rapid assessment of patients' constitution. Unlike during diagnosis, the applications of ML in before or after diagnosis stages are less common. A possible reason is that the early research focused on using clinical data to train ML models and assist in diagnosis. As attention to pre-diagnosis questionnaires and post-treatment follow-ups increases, the use of ML is expected to grow in these areas.

The applications of ML in TM diagnosis primarily focus on the during diagnosis stage, where ML techniques serve as an auxiliary tool during the patient examination and treatment process. Many works utilize the classification capabilities of ML models to assist in diagnosis. Ref. 70 applied feature selection and syndrome classification on chronic hepatitis B patient records to construct a TCM diagnosis model for syndrome differentiation formulas. Ref. 71 used a Support Vector Machine (SVM) model to classify lip images to assist in TCM lip diagnosis. Ref. 72 developed SVM and K-Nearest Neighbors (KNN) models based on the relationship between syndromes and symptoms from clinical cases to classify patients with hypertension. Ref. 73,75,80 utilized ML classification models to recommend prescriptions based on the symptoms presented by patients, assisting in diagnosis and treatment. Ref. 78 proposed an SVM classification model using unstructured pulse signal data for pulse signal classification, which is valuable for objectifying pulse diagnosis. Ref. 79 developed Random Forest (RF) and DT models for classifying TCM constitutions based on volunteer test data. Apart from classification tasks, Ref. 76 tracked the treatment of thromboembolic patients using warfarin and developed a multivariate nonlinear regression ML ensemble model for quantitative dosage prediction, minimizing risks of incorrect dosing. Additionally, clustering models of ML can also assist in diagnosis. Ref. 74,77 further proposed their improved algorithms based on K-means clustering models, applying them to the clustering analysis of Chinese medicine prescriptions and supporting clinical decision-making.

4.2. ML applications in research

Apart from diagnosis, ML is also widely applied in TM research, primarily in the two areas: drug research and data analysis. In drug research, ML can be used not only to identify herbs at a macro level but also to recognize target compounds within herbs at a micro level. Ref. 82,84,89 used spectral data of Chinese medicinal herbs to develop ML classification models for efficient identification and differentiation, advancing herbal drug discovery. Ref. 83 employed a homemade E-nose system to collect ginseng data and developed a high-accuracy ginseng classification model using ML ensemble learning. Ref. 86 and87 collected herbal medicines from prescriptions and proposed improved K-means clustering models to more accurately identify herbs of the same category. Ref. 88 developed ML classification models to rapidly identify TCM with activity against E. coli and/or S. aureus. from a large antibacterial herb dataset. At the micro level, Ref. 81 used compound fingerprint features to create Naive Bayesian (NB) classification and recursive partitioning models, facilitating the discovery of drug-like molecules. Ref. 85 and91 utilized molecular feature data to identify target compounds, focusing on neuroprotective compounds from XiaoXuMing Decoction (XXMD, a TCM prescription) and hepatotoxic compounds from TCM and Western medicine combinations. Ref. 90 used HepG2 cells to investigate the hepatotoxicity risks of natural products derived from TCM, developing an SVM model to quantitatively predict hepatotoxicity and recommend dosing regimens. Ref. 92 developed ML classification models to screen combinations of Western medicine and TCM for anti-RA treatment from small molecule drug datasets, providing a reference for clinical treatment of RA with integrated Western medicine and TCM.

Utilizing ML models for data analysis to enhance the interpretability of TM in practical applications is also a valuable direction. Ref. 93 used a DT model to study the relationship between posthepatitic cirrhosis syndromes and symptoms or signs, extracting differentiation rules and identifying key symptoms or signs., Ref. 94,101 and102 adopted RF models for factor analysis to identify the risk factors of osteoporosis, the key aspects of TCM diagnosis and treatment for insomnia, and the most frequently used drugs for treating acute pancreatitis, respectively. Ref. 95 and99 focused on herbal combinations and used ML models to analyze the cardiotoxicity of TCM compatibility and incompatible herb pairs, respectively. Ref. 96 used an SVM model to identify articles on TM from biomedical literature, facilitating research trend analysis in TM. Ref. 97 proposed a chromatographic strategy combined with an SVM model for quality testing of KuDieZi injection, advancing the quality analysis of TCM injections. Ref. 98 suggested a Genetic Algorithm-Backward Propagation Neural Network (GA-BPNN) model to analyze and establish a mapping between process parameters and extraction yield in the bioactive ingredient extraction from Chinese herbal medicine, further optimizing the extraction parameters. Ref. 100 developed ML models to analyze target meridians of herbs and their active components in traditional Chinese herbal medicine.

Research on ML for structured knowledge is relatively limited, as there is a growing preference for more complex and powerful models. For instance, DL models are often used for Named Entity Recognition (NER) to construct knowledge graphs, or LLMs are applied to extract relational structures from textual data. These approaches will be discussed in detail in subsequent sections.

5. DL on traditional medicine

With the rapid development of artificial intelligence, Deep Learning (DL), a key branch of machine learning, offers new avenues for exploring TM. It allows automatic learning of complex patterns and relationships through multi-level feature extraction, nonlinear transformation, and end-to-end learning. The applications of DL in TM are diverse and extensive. The ongoing development of DL models, like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Graph Neural Networks (GNN), provides innovative approaches for researchers, improving the efficiency of medical data processing and patient treatment, and advancing the medical field. As shown in Table 3, this section systematically reviews DL in TM and introduces representative works in terms of diagnosis and research.

Table 3
Overview and statistics of DL Applications in Traditional Medicine.

Application	Description	Model	Data (Size)	Publications (Year)
During diagnosis	Quantifiable TCM pulse diagnostic model	ANN	Volunteer subjects body data (260)	[103] (2012)
	Non-Invasive Multi-Disease Classification via Facial Image Analysis	CNN	Facial images (1181)	[104] (2018)
	Classification of prescription and Medicine SE Prediction	ANN	Safe prescriptions (150)	[105] (2019)
	Classification of Traditional Chinese Medicine Cases	CNN, LSTM	Medical cases (40000)	[106] (2019)
	Traditional Medicine Prescription Generation	LSTM	Insomnia data (6428)	[107] (2019)
	Detect and classify hypertension	CNN	Tongue image (4000)	[108] (2020)
	TCM case text classification	RNN, CNN	Medical record data (12053)	[109] (2021)
	Chinese Medical Text Classification	GNN	Medical periodicals (14482)	[110] (2021)
	TCM prescription diagnosis and recommendation system	CNN	EHR (20000)	[111] (2021)
	Herbal prescription multi-label classification	LSTM	Records of prescription syndrome (385)	[112] (2021)
	prediction and classification of Chinese prescriptions	GNN	Angina targets (569)	[113] (2023)
	Integrated image and location analysis for wound classification	CNN	Wound images (1268)	[114] (2024)
	Exploring hepatic fibrosis screening via tongue images	CNN	Tongue image (741)	[115] (2024)
Drug research	Classification of Chinese herbal medicine	CNN	Chinese herbal image (1700)	[116] (2019)
	Leaf Image Classification for Bangladeshi Medicinal Plant Recognition	CNN	Medicinal plant images (37693)	[117] (2020)
	A mobile application for the recognition of medicinal plants	CNN	Medicinal plant images (1000)	[118] (2021)
	Attentive Herb Recommendation	GNN	Drug prescriptions (26360)	[119] (2021)
	Recognize Bangladeshi Ordinary Medicinal Plants	CNN	Medicine leaves images (2000)	[120] (2021)
	Herbal 2D-COS spectral image classification	CNN	Herbal samples (429)	[121] (2022)
	Herbal efficacy multi-label classification	CNN, SVM	Unlabeled formulas (31114)	[122] (2023)
	Detecting True Medicinal Leaves	CNN, SVM, ANN	Leaf images (9000)	[123] (2023)
	Classification of TCM	CNN	Chinese medicine image (7853)	[124] (2023)
	Traditional Vietnamese Herbal Medicine Image Recognition	ANN, CNN, SVM	Images of herbs (10000)	[125] (2023)
Structured knowledge	Named Entity Recognition of TCM Patents	LSTM	Chinese medicine patent abstract texts (2000)	[126] (2021)
	Traditional Chinese medicine entity relation extraction	CNN, KNN, SVM	Samples of herbal chemicals (4732)	[127] (2022)
	Knowledge Graph in TCM Diagnosis and Treatment of Viral Hepatitis B	LSTM	Medical records of patients (2573)	[128] (2022)
	A Strategy for TCM Rare Disease Syndrome Differentiation	CNN, RNN	TCM clinical records (1123)	[129] (2023)
	Complex Network Modeling Approach to Explore TM	GNN	compound-target interactions(1900)	[130] (2022)

5.1. DL applications in diagnosis

The diagnosis of TM is not only the starting point of clinical treatment but also the key factor in determining treatment methods and outcomes. Diagnostic results directly influence subsequent treatment choices, making the application of DL techniques in this stage central. TM's application of DL is primarily focused on the diagnostic stage because diagnosis relies on quantifiable data (such as tongue patterns, pulse, complexion, etc.), which are well-suited for image recognition and pattern analysis using DL. In contrast, the before-diagnosis and after-diagnosis stages lack long-term tracking data (such as lifestyle habits, recovery processes, etc.), and the data often involves patient privacy, making it difficult to directly apply DL techniques. Consequently, it is relatively hard to find the works for before and after diagnosis, whereas we find many works belong to the during-diagnosis stage.

The diagnostic process in TM largely relies on the experience and judgment of practitioners, making it difficult to describe in a scientifically quantitative manner. Therefore, practitioners leverage DL models to process patients' facial features, irises, tongue coatings, and other indicators, helping the development of standardized, objective, and quantifiable diagnoses. For example, the Four Diagnostic Methods—inspection, auscultation, inquiry, and palpation—are used in TCM, where practitioners observe the five organs (heart, liver, spleen, lung, kidney) through the appearance of the five facial features (mouth, ears, eyes, nose, tongue). Among them, according to the application of the DL of the tongue, the appearance of the tongue reflected the organs and the severity of the disease such as Ref. 108,120. Ref. 121 developed a CNN model to evaluate the effectiveness of tongue image analysis for the detection of liver fibrosis. Ref. 111 adopted DL in TCM visual examination and iridology to predict organ dysfunction and provide health advice. Ref. 104 extracted specific facial regions to create a non-invasive multi-disease classifier using CNN. Ref. 103 introduced a TCM pulse diagnostic model to quantify pulse diagnosis and differentiate essential hypertension from normotensive back propagation. Ref. 116 developed a multimodal CNN model that classified wounds into diabetes, pressure, surgery, and venous ulcers, improving the precision of diagnosis.

On the other hand, medical records are often continuous records of prescribed medications, typically containing a large number of TM names, herbal combinations, and usage patterns. Regarding medical records, DL focuses on text analysis, efficiently processing complex prescription texts through case classification to explore the relationships between different prescriptions and medical conditions. Ref. 107 collected and classified prescriptions from ancient Chinese medicine classics to assess whether prescriptions would have side effects according to cold and heat. Ref. 105,[106],113 captured the characteristics of pure text data in prescriptions to achieve efficient retrieval of medical information. Ref. 112 employed the LSTM model to explore the efficacy of formulation included in law data. Ref. 103 designed the TM recommendation system to carry out the prescription generation. Furthermore, GNN is a type of deep learning models for graph data, and it can help us better understand the association and role of different entities in the TM text data. Ref. 130 used GNN to predict unknown compound marker interactions. Ref. 113 connected Chinese medicine prescriptions with images to create a herbal recommendation system. Ref. 110 improved model performance by integrating GNN with knowledge graph for medical text classification.

5.2. DL applications in research

The application of DL in the research area focuses primarily on drug discovery, followed by structured knowledge, with relatively less emphasis on data analysis. As for the reasons behind the limited application of data analysis, we hypothesize that TM emphasizes a holistic approach and dialectical thinking, viewing diseases as dynamic systems rather than isolated individuals. This approach stresses comprehensive analysis over quantitative data evaluation. The quantitative indicators are less frequently used in the domain of TM in the past, and accordingly there is a lack of standardized criteria and sufficient labeled data, which limits the application of DL in data analysis.

In TM, herbs are an essential component of treatment, and their physical characteristics (such as appearance, color, shape, etc.) play a crucial role in traditional drug research. DL employs image recognition techniques to efficiently and accurately process these complex visual data, making it as the core technique for herb identification and database construction. We divide the application of DL in drug research into two categories: one is helping practitioners identify herbs, extract their features, and map them to structured data in herb databases; another one is assisting physicians to acquire medical knowledge, establish the relationship between herbs and symptoms, and achieve accurate classification of herbal symptoms.

In the first category of drug research, Ref. 115,124,125 developed automatic systems for classifying medicinal plants to help people quickly identify useful herbs. Ref. 127 developed a mobile application that identifies 70 different medicinal plants to help non-botanists recognize the medicinal plants they want. Ref. 119 classified the correct drug leaves from other leaves of similar appearance with a deep convolutional neural network. Ref. 130 proposed a CNN-based method to classify traditional Vietnamese herbal images. Ref. 117 developed an automated system using a CNN to classify 10 medicinal plants in Bangladesh. Ref. 118 created a mobile app using the CNN model to identify 70 medicinal plants in Mauritius.

In the second category of drug research, Ref. 122 proposed a Region-based Convolutional Neural Network (RCNN) strategy to promote 2D-COS spectral images in medicinal plants and related fields of applications for the quality control of herbal medicine. TCM doctors use the synergistic effect of active substances to enhance efficacy and reduce side effects. Ref. 123 proposed a deep learning-based separation feature extraction method to improve the extraction of features and the prediction of the efficacy of herbal medicines and to find anon-linear relationship between the formula and the efficacy of herbal medicines.

A large amount of information in TM, such as herbal formulas, disease descriptions, and treatment methods, often exists in unstructured form. DL techniques can automatically extract and structure them from complex data, significantly enhancing the efficiency of models in medical applications. Ref. 126 used an LSTM model to identify herbs, diseases, and treatments in TCM patents. This method improved entity recognition without manual feature engineering. Ref. 127 employed a CNN model with a segment attention mechanism to extract herbal relations from PubMed. The technique improved the classification accuracy of the relation in herbal datasets. Ref. 129 introduced a Two-Level Data Association (TLDA) model to address rare diseases in TCM by enriching limited Electronic Medical Record (EMR) data. TLDA achieved superior results in diagnosing with small datasets. Ref. 128 built a TCM knowledge graph for hepatitis B using real medical data. This graph enhanced a Q&A system for better TCM-based diagnosis and patient care.

6. LLM on traditional medicine

As an emerging research direction of AI in recent years, Large Language Models (LLM or LLMs) have yielded many exciting outcomes when combined with TM. From the computational viewpoint, since LLMs excel in generative tasks within natural language tasks, making this a primary focus area. However, LLMs also have the ability to handle tasks such as classification. From a medical point of view, the tasks that LLMs can address are diverse. Medical diagnostic tasks are mainly centered on the generation of prescriptions or prediction of symptoms, while research tasks primarily involve structured knowledge and data analysis based on knowledge extraction. Unlike traditional ML techniques, LLMs offer a unique comprehensive model that can solve multiple problems and determine needs through multiple interactions. Such models perform excellently in both diagnostic and research tasks, making them one of the key directions for future development. This section covers three subsections based on TM applications: diagnosis, research, and general models, as shown in Table 4. It should be noted that the number of parameters is an important characteristic of LLMs, and the estimated number of model parameters is denoted by ~ if no such information is provided in the original papers.

6.1. LLM applications in diagnosis

Regarding LLMs for diagnostic purposes, Ref. 131 completed the diagnostic assistance for chest impediments, while Ref. 136 focused on diagnostic tasks related to digestive issues. The latter also emphasized the importance of fine-tuning techniques, which allow for improving performance in specific domains. Similarly, regarding comprehensive tasks, Ref. 133–135 tried to fine-tune pretrained LLMs so as to accomplish prescription generation tasks in the medical field. Notably, Ref. 133 highlighted the enhancement of fine-tuning when the sample size is limited, whereas Ref. 134 established a knowledge graph to enhance the model's knowledge base. Additionally, Ref. 132 conducted symptom-based condition diagnoses.

6.2. LLM applications in research

The most significant application of LLMs in research is its powerful knowledge extraction capability, allowing them to handle both structured and unstructured knowledge. Named Entity Recognition (NER) assists researchers extract information from various sources including medical cases,[140] social media,141 and ancient texts.146 Furthermore, Ref. 145,146 enabled the construction of structured knowledge through NER. Ref. 139 retrieved relevant content from vast amounts of information for investigation, while Ref. 144 extracted these contents for constructing a reminder system. From a technical standpoint, transfer learning[143] or semi-supervised learning[142] can be used to address the problem of insufficient data. Regarding drug research, NER can also assist the relationship extraction of drugs, which helps in researching drug side effects.[137] Other drug-related research can combine knowledge graphs and fine-tuned models to establish dialogue models to support the drug-related research.[138] Besides, it is also feasible to directly fine-tune LLMs for classification and data analysis. Ref. 147,148 classified medical cases, Ref. 149 classified medical literature, and Ref. 150,151 classified acupuncture points.

6.3. LLM applications for general purpose

LLMs are capable of serving as general-purpose models for handling multiple tasks, such as single-choice questions, symptom prediction, prescription recommendations, and basic QA. This is possible because TM corpora are versatile, encompassing sources such as ancient books, medical records, QA data, and structured data. These comprehensive corpora have enabled the development of mature and robust models, offering diverse options for TM applications.

For TM-specialized models, they are trained on various high quality TM datasets. Ref. 157 developed the Qibo dataset, approximately 2 GB in size, and employed continuous pre-training to mitigate over confidence issues. Meanwhile, Ref. 154 utilized 20 GB of on-line sources combined with the TCM-EXAM and TCM HER datasets to form the TCM-Corpus-1B dataset, applying domain-specific corpora for efficient

Table 4
Overview and statistics of LLM Applications in Traditional Medicine.

Application	Description	Dataset	Model (Parameters)	Publications (Year)
During diagnosis	Classification of chest impediment medical records	Records	BERT (~110M)	[131] (2019)
	A finetuned TCM model to predict final diagnosis	Records	BERT (~110M)	[132] (2021)
	Prescription generation with small quantity dataset	Records, ancient books	BERT (~110M)	[133] (2022)
	A finetuned TCM model to generate prescription with KG	Records, structure data	BERT (~110M)	[134] (2023)
	TCM prescription generation with GPT-4	Records, textbooks	GPT-4 (1800B)	[135] (2023)
	TCM-FTP, finetuned prescription generation model to solve digestive problems	Records	LLaMA (7B)	[136] (2024)
Drug research	Extract adverse drug reaction-related information	Records	BioBERT (110M)	[137] (2021)
	Medication guidance, drug interaction prediction, adverse reaction prediction	Structure data	ChatGLM (6.2B)	[138] (2024)
Structured knowledge	Entity and relation extraction in medical text	Records	BERT (110M)	[139] (2019)
	NER in electronic medical records	Records	BERT (110M)	[140] (2020)
	Local drug name recognition, supporting identification of drug usage trends	Online sources	BERT (110M)	[141] (2020)
	TCM terminology recognition with semi-supervised fine-tuning	Records	BERT (~100M)	[142] (2020)
	TCM terminology recognition using transfer learning	Online sources	BERT (~100M)	[143] (2021)
	NER in TCM corpus to extract drugs to build a reminder system	Online sources	RoBERTa (125M)	[144] (2022)
	TCM knowledge structuring and integration	Online sources	IFlytekSpark (1000B)	[145] (2024)
	TCM KG construction using NER on an ancient book	Ancient books	BERT (~100M)	[146] (2024)
Data analysis	Classification of TCM clinical records	Records	BERT (110M)	[147] (2019)
	Classification and information extraction of TCM medical records	Records	BERT (~100M)	[148] (2020)
	Medical research literature classification	Abstracts	BERT (110M)	[149] (2020)
	Acupuncture point classification and recommendation	Structure data	BERT (110M)	[150] (2021)
	Relation extraction and standardization of acupuncture point locations	Structure data	GPT-3.5 (175B)	[151] (2024)
General purpose	EpidemicCHAT, a general model of TM epidemic prevention and treatment	Records, structure data, ancient books, QA data	ChatGLM (13B)	[152] (2023)
	Zhongjing, a TCM model supporting multiple tasks	QA data	LLama2 (7B)	[153] (2023)
	TCM-GPT, a TCM model supporting multiple tasks	Records, QA data, online sources	BLOOM(7B)	[154] (2024)
	Lingdan, LLM models with specific models supporting multiple tasks using CoT	Structure data, records, ancient books, textbooks	Baichuan2 (13B)	[155] (2024)
	MedChatZH, a general model specializing in QA ability	QA data, textbooks	Baichuan (7B)	[156] (2024)
	Qibo, a TCM model supporting multiple tasks based on LLaMA	Structure data, records, ancient books, textbooks	LLaMA (7B/13B)	[157] (2024)
	BenTsao model, a series of models fine-tuned on LLaMA models	Structure data	LLaMA (7B)	[158] (2023)
	DoctorGLM, a medicine model fine-tuned on ChatGLM	Structure data	ChatGLM (6B)	[159] (2023)
	Palm2, one of the biggest medical general models published by Google	/	Palm2 (340B)	[160] (2023)
	GPT-4, one of the biggest general models	/	GPT-4 (1800B)	51 (2023)

pre-training. Similarly, Ref. 155 designed various models and pre-trained them on TCMPT, followed by fine-tuning on the TCPM QA dataset and SSHPR to address different tasks. Ref. 153,156 and 152 focused on QA tasks. Specifically, Ref. 153 enhanced its capabilities using the ShenNong TCM Dataset, containing 112,564 QA pairs. Ref. 156 used over 1000 textbooks in the pre-training stage, followed by fine-tuning on 764,000 conversation datasets to enhance its performance. In contrast, Ref. 152 compiled 6.9 million instruction data, combining classical texts and structured knowledge like knowledge graphs, to augment the training data for improved task performance.

For non-specialized models, these models, while not trained on dedicated TM datasets, still demonstrate the ability to solve TM-related problems. For example, Ref. 158,159 as Chinese general medical models, achieved good results in comparison experiments in Ref. 157. In addition, both the English general medical model[160] and general models like Ref. 51 possess the additional capability to address TM issues. This is because open-world corpora contain some TM-related data, allowing models not specifically trained on TM to also have.

6.4. Further discussion on LLMs

To provide further insights and practices of current LLMs applied to TM, we selected three different types of questions, covering medical diagnosis, single-choice questions, and knowledge-based queries; we empirically investigated five different LLMs, including one general model, two general medical models, and two TM models to showcase the current strengths and weaknesses. As shown in Tables 5–7 in Appendix, the overall performance of LLMs is impressive. First, both TM models accurately provided the corresponding prescription name for the symptom in the medical diagnosis task. Second, the general medical models demonstrated strong logical reasoning and gave the correct answer to the single-choice question. Third, all models including the general model performed exceptionally well in knowledge-based query. However, there is still significant room for improvement.

On the one hand, most specialized models can only perform prescription recommendation tasks, with weaker capabilities in other knowledge-based tasks. Key areas for advancement include constructing more structured datasets, employing fine-tuning or prompt engineering to enhance model performance. Further, as shown in Table 5, which includes the responses of Zhi Hui Ling Yi, the TM responses of general models are often interfered with by non-TM approaches rather than being effectively integrated. Specifically, general models and general medical models tend to over-rely on information provided by non-TM diagnostic methods, leading to a reluctance to prescribe TM treatments. Therefore, developing systems to compare TM with non-traditional approaches, and leveraging TM methods and philosophies to support non-TM practices in patient diagnosis and treatment, are promising steps towards improving the synergy between these systems and enhancing patient care.

On the other hand, the accuracy of tasks performed by most LLMs is not yet at an optimal level, including hallucinations and inaccuracies commonly seen in large models, as well as the issues in comprehension that arise in small models due to their limited scale. These situations highlight substantial room for improvement in performance metrics. Besides, LLMs face unresolved challenges, particularly ethical concerns related to explainability, security, and reliability, which stem from the inherent opacity of their computations. Addressing these issues will require enhancing interpretability through methods such as Chain-of-Thought (CoT) reasoning or integrating knowledge graphs to validate outputs. Additionally, refining LLMs' logical reasoning capabilities and fostering closer collaboration with healthcare professionals are critical steps toward more effective and responsible use of these models in the medical field.

In terms of application, some commercial APIs have already been developed, such as Zhi Hui Ling Yi[5] from Baidu. These projects offer comprehensive services such as diagnostic assistance and one-stop medical record management. However, many techniques of LLM applications still require significant advancements for broader adoption and deeper deployment. One important area is the comprehensive integration of processes before and after diagnosis. For example, LLMs could provide lifestyle or dietary recommendations based on the patient's medical history, in addition to simple prescribing medication or self-service inquiries. Another key direction is offering more support to doctors during the diagnostic process. This includes tasks like symptom classification before diagnosis, generating more constructive treatment suggestions during diagnosis, and providing interpretable treatment plans to help doctors make better-informed decisions.

7. Discussions

7.1. Data insufficiency, modality, and privacy

Data is essential for both medical diagnosis and research, but vast amounts of data are required to train models effectively. These models learn the underlying patterns from data and aim to adapting to unseen inputs. As a result, the progress of AI in TM is highly related to the growth of relevant data. However, the data issues also present a critical challenge that needs to be addressed including the lack of database and imbalance in multi-modal and cross-cultural development.

First, the amount of structured TM data available for training LLMs is insufficient. Although ancient books contain a wealth of textual contents, most of them have not yet been systematically structured. If more ancient data could be digitized and structured, it would greatly enhance the learning efficiency of LLMs and mitigate the current fragmentation of independent datasets. Second, while current LLMs focus primarily on text processing tasks, multi-modal tasks, such as incorporating tongue diagnosis images, are also worth exploring to better align with the holistic characteristic of TM. Third, although TM databases are gradually expanding, data resources for non-TCM traditional medicine such as Indian TM remain relatively scarce.

These challenges are largely due to medical privacy concerns and technical limitations. On one hand, patient privacy and data security restrict the open sharing and structuring of medical data. For example, the EU General Data Protection Regulation (GDPR)[161] imposes strict limitations on the use of personal data, making it challenging to use medical records as training data for AI models. On the other hand, the processing of multi-modal data, such as tongue diagnosis images, requires advanced techniques and computational resources, which are currently insufficient for widespread application. Therefore, collaborative large-scale dataset projects across multiple hospitals are essential. Such initiatives not only address data and resource scarcity but also enable consistent anonymization of patient information, enhancing data security. Additionally, they can provide diverse and high-quality data sources, which support more comprehensive applicable research outcomes, fostering advancements across varied medical fields.

7.2. From unimodal to multimodal models

Considering the perspective of AI models, the advancement of multimodal AI models is also crucial in the field of TM. TM, e.g., TCM, emphasizes holistic diagnosis, requiring a comprehensive assessment through visual inspection, auscultation, inquiry, and palpation to provide accurate treatment. Likewise, AI models also require integrating diverse data modalities to make well-informed recommendations. Currently, most AI models for TM focus on a single data type, such as text, images, sequence signals, or one-hot encoding. However, emerging methods like Statistically Enhanced Learning[162] show promising potential for addressing the challenges of multi-modal data learning in TM. Ref. 163 has investigated DL-based multi-modal fusion methods for various data types, but these models encounter challenges such as generalizability and interpretability. Additionally, as previously mentioned, data collection remains a critical aspect. Few researchers fully recognize the importance of collecting diverse data for multi-modal models, which restricts the development of multi-modal AI in TM.

In recent years, the general purpose LLMs for medicine have gained popularity, with multi-modal LLMs emerging as a promising research direction.[164–166] To our best knowledge, no multi-modal LLMs has yet been specifically designed for TM. However, with advancements in multi-modal AI and growing interest in TM, significant breakthroughs in this direction are likely on the horizon.

7.3. Preventive healthcare

The preventive healthcare, which could reduce disease risk through early intervention, is vital for enhancing patient outcomes and safeguarding public health. For instance, the "Healthy China 2030″ initiative[6] campaigns to prevent disease and promote health.[167,168] Similarly, the U.S. "Healthy People" program[7] focuses on setting health goals, enhancing public health services, and promoting disease prevention.[169,170] These policies underscore the need of disease prevention and health management, and encourage to adopt new technologies like AI in healthcare.

As for TM, the applications of AI in the stages of before and after diagnosis are relatively limited due to data scarcity and accessibility issues. Unlike the abundant data collected during diagnosis, the data for before and after diagnosis, are often fragmented and neglected, particularly in disease prevention and health management. But these data are vital for preventive healthcare. As a result, further efforts should be dedicated to improve data collection and sharing mechanisms to better support preventive healthcare. By prioritizing prevention, TM evolves from a passive diagnosis to proactive diagnosis, thereby aligning with global health trends.

7.4. Trustworthiness

The trustworthiness of AI has recently attracted much attention[171,172] and has also been suggested by several governments and organizations such as the Recommendation on the Ethics of Artificial Intelligence by UNESCO[173] and the Ethical Norms for New Generation Artificial Intelligence by China.[174] Notably, the trustworthy AI is particularly important while employing it to medical applications comparing to other applications.[175–177] However, the existing AITM works mostly aim to adapt AI techniques for suitable TM problems and attempt to improve the performance of AI models, while these works

[5] https://01.baidu.com/

[6] https://www.who.int/teams/health-promotion/enhanced-wellbeing/ninth-global-conference/healthy-china.

[7] https://www.cdc.gov/nchs/healthy_people/index.htm.

rarely consider the trustworthiness of AI models. To enhance the credibility of AI models, it is imperative to invest more efforts in the direction of trustworthiness research for AITM, which involves delving into aspects such as model explainability, interpretability, transparency, reliability, and robustness. Due to the concerns about trustworthiness and ethics, AITM is still regarded as a tool for human TM practitioners. Its development should focus on becoming more effective computer assistants rather than aiming to replace clinicians and doctors.

8. Conclusion

This paper presented a comprehensive review of the recent advancements of AI in the context of TM. The review not only encompassed the recent machine learning and deep learning techniques applied to TM, but also delved into the emerging technique of large language models for TM. To be more specific, the research works of AITM were first categorized into the three recent AI techniques according to the parameter scale and the model architecture. For each AI technique, research works were further grouped based on their applications for diagnosis (before, during, after) or research (drug research, structured knowledge, data analysis). And for each research work, we summarized its key information in tables and described its main contents. On the basis of these research works, we discussed the challenges and highlighted the promising future directions. Apart from these research works and challenges, we also summarized the existing review papers related to AI in both traditional and non-traditional medicine. We hope this review can present the overall picture of existing AITM research works and provide some insights for the future exploration. Besides, the interdisciplinary collaboration and innovation are highly recommended to further propel AITM forward and amplify its impact on human health.

Author contributions

Chengbin Hou and William Cheng-Chung Chu conceptualized this review and designed the main contents; Chengbin Hou wrote Sections 1 and 3, and optimized the whole manuscript; Yanzhuo Gao wrote Section 5, and refined Section 2, all figures and format; Xinyu Lin wrote Section 4 and enhanced all tables; Jinchao Wu wrote Section 6 and conducted the LLM experiments in the appendix; Ning Li prepared the initial draft of Section 2; Chengbin Hou, Yanzhuo Gao, Xinyu Lin, and Jinchao Wu jointly wrote Section 7 and the conclusion; William Cheng-Chung Chu and Hairong Lv provided supervision and valuable comments, and revised the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the anonymous reviewers for their valuable advice. This work was funded by the National Key R&D Program of China (Grant No. 2022YFF1202400) and the Henan Province Medical Science and Technology Research Plan Major Project jointly built by the province and the ministry (No. SBGJ202401001).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jtcme.2025.02.009.

References

1. W. H. Organization. *WHO Traditional Medicine Strategy: 2014-2023*. World Health Organization; 2013.
2. Patwardhan B, Partwardhan A. *Traditional Medicine: Modern Approach for Affordable Global Health*. Switzerland: World Health Organization; 2005.
3. Payyappallimana U. Role of traditional medicine in primary health care. *Yokohama J Soc Sci*. 2010;14:57–75.
4. da Rosa C. Traditional medicine and complementary/alternative medicine in primary health care: the brazilian experience. In: *Primary Care at a Glance-Hot Topics and New Insights*. IntechOpen; 2012.
5. Dobos G, Tao I. The model of western integrative medicine: the role of Chinese medicine. *Chin J Integr Med*. 2011;17:11–20.
6. Russell SJ, Norvig P. *Artificial Intelligence: A Modern Approach*. Pearson; 2016.
7. Li W, Ge X, Liu S, Xu L, Zhai X, Yu L. Opportunities and challenges of traditional Chinese medicine doctors in the era of artificial intelligence. *Front Med*. 2024;10, 1336175.
8. Wang S, Hou Y, Li X, Meng X, Zhang Y, Wang X. Practical implementation of artificial intelligence-based deep learning and cloud computing on the application of traditional medicine and western medicine in the diagnosis and treatment of rheumatoid arthritis. *Front Pharmacol*. 2021;12, 765435.
9. Chu H, Moon S, Park J, Bak S, Ko Y, Youn B-Y. The use of artificial intelligence in complementary and alternative medicine: a systematic scoping review. *Front Pharmacol*. 2022;13, 826044.
10. Khan SR, Al Rijjal D, Piro A, Wheeler MB. Integration of AI and traditional medicine in drug discovery. *Drug Discov Today*. 2021;26:982–992.
11. Arji G, Safdari R, Rezaeizadeh H, Abbassian A, Mokhtaran M, Ayati MH. A systematic literature review and classification of knowledge discovery in traditional medicine. *Comput Methods Progr Biomed*. 2019;168:39–57.
12. Zhou E, Shen Q, Hou Y. Integrating artificial intelligence into the modernization of traditional Chinese medicine industry: a review. *Front Pharmacol*. 2024;15, 1181183.
13. Wang Y, Shi X, Efferth T, Shang D. Artificial intelligence-directed acupuncture: a review. *Chin Med*. 2022;17:80.
14. Zhang P, Zhang D, Zhou W, et al. Network pharmacology: towards the artificial intelligence-based precision traditional Chinese medicine. *Briefings Bioinf*. 2024;25, bbad518.
15. Chu X, Sun B, Huang Q, Peng S, Zhou Y, Zhang Y. Quantitative knowledge presentation models of traditional Chinese medicine (tcm): a review. *Artif Intell Med*. 2020;103, 101810.
16. Ng JY, Cramer H, Lee MS, Moher D. Traditional, complementary, and integrative medicine and artificial intelligence: novel opportunities in healthcare. *Integr Med Res*. 2024, 101024.
17. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. *Nat Med*. 2022;28:31–38.
18. Hamet P, Tremblay J. Artificial intelligence in medicine. *Metabolism*. 2017;69: S36–S40.
19. Ahmad Z, Rahim S, Zubair M, Abdul-Ghafar J. Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. a comprehensive review. *Diagn Pathol*. 2021;16:1–16.
20. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol*. 2019;28:73–81.
21. Becker A. Artificial intelligence in medicine: what is it doing for us today? *Health Pol Technol*. 2019;8:198–205.
22. Johnson KB, Wei W-Q, Weeraratne D, et al. Precision medicine, AI, and the future of personalized health care. *Clin Transl Sci*. 2021;14:86–93.
23. Wang R, Pan W, Jin L, et al. Artificial intelligence in reproductive medicine. *Reproduction*. 2019;158:R139–R154.
24. Habuza T, Navaz AN, Hashim F, et al. AI applications in robotics, diagnostic image analysis and precision medicine: current limitations, future trends, guidelines on CAD systems for medicine. *Inform Med Unlocked*. 2021;24, 100596.
25. Nensa F, Demircioglu A, Rischpler C. Artificial intelligence in nuclear medicine. *J Nucl Med*. 2019;60:29S–37S.
26. Panayides AS, Amini A, Filipovic ND, et al. AI in medical imaging informatics: current challenges and future directions. *IEEE J Biomed Health Inf*. 2020;24: 1837–1857.
27. Wang F, Preininger A. AI in health: state of the art, challenges, and future directions. *Yearbook Med Inf*. 2019;28:16–26.
28. Kirubarajan A, Taher A, Khan S, Masood S. Artificial intelligence in emergency medicine: a scoping review. *J Am College Emerg Phys Open*. 2020;1:1691–1702.
29. Hamamoto R, Suvarna K, Yamada M, et al. Application of artificial intelligence technology in oncology: towards the establishment of precision medicine. *Cancers*. 2020;12:3532.
30. Toh TS, Dondelinger F, Wang D. Looking beyond the hype: applied AI and machine learning in translational medicine. *EBioMedicine*. 2019;47:607–615.
31. Hirani R, Noruzi K, Khuram H, et al. Artificial intelligence and healthcare: a journey through history, present innovations, and future possibilities. *Life*. 2024; 14:557.
32. Subramanian M, Wojtusciszyn A, Favre L, et al. Precision medicine in the era of artificial intelligence: implications in chronic disease management. *J Transl Med*. 2020;18:1–12.
33. Quazi S. Artificial intelligence and machine learning in precision and genomic medicine. *Med Oncol*. 2022;39:120.

34. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25:44–56.
35. Zhang P, Kamel Boulos MN. Generative AI in medicine and healthcare: promises, opportunities and challenges. *Future Internet*. 2023;15:286.
36. Alam SF, Gonzalez Suarez ML. Transforming healthcare: the ai revolution in the comprehensive care of hypertension. *Clin Pract*. 2024;14:1357–1374.
37. Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. *Gastrointest Endosc*. 2020;92:807–812.
38. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. *Nat Med*. 2019;25:30–36.
39. Lu M, Yin J, Zhu Q, et al. Artificial intelligence in pharmaceutical sciences. *Engineering*. 2023;27:37–69.
40. Kolluri S, Lin J, Liu R, Zhang Y, Zhang W. Machine learning and artificial intelligence in pharmaceutical research and development: a review. *AAPS J*. 2022; 24:1–10.
41. Haug CJ, Drazen JM. Artificial intelligence and machine learning in clinical medicine, 2023. *N Engl J Med*. 2023;388:1201–1208.
42. Selvaraj C, Chandra I, Singh SK. Artificial intelligence and machine learning approaches for drug design: challenges and opportunities for the pharmaceutical industries. *Mol Divers*. 2021:1–21.
43. Zhang Y. *An In-Depth Summary of Recent Artificial Intelligence Applications in Drug Design*. 2021. arXiv preprint arXiv:2110.05478.
44. Deng J, Yang Z, Ojima I, Samaras D, Wang F. Artificial intelligence in drug discovery: applications and techniques. *Briefings Bioinf*. 2022;23, bbab430.
45. Jiménez-Luna J, Grisoni F, Weskamp N, Schneider G. Artificial intelligence in drug discovery: recent advances and future perspectives. *Expet Opin Drug Discov*. 2021; 16:949–959.
46. Morley J, Machado CC, Burr C, et al. The ethics of AI in healthcare: a mapping review. *Soc Sci Med*. 2020;260, 113172.
47. Sharma A, Al-Haidose A, Al-Asmakh M, Abdallah AM. Integrating artificial intelligence into biomedical science curricula: advancing healthcare education. *Clin Pract*. 2024;14:1391–1403.
48. Ryzhova K, Yumashev A, Klimova M. Artificial intelligence in the diagnosis of diseases of various origins. *J Complemen Med Res*. 2023;14:199.
49. Liu R, Rong Y, Peng Z. A review of medical artificial intelligence. *Global Health J*. 2020;4:42–45.
50. Devlin J, Chang M-W, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics; 2019:4171–4186.
51. Achiam J, Adler S, Agarwal S, et al. *Gpt-4 Technical Report*. 2023. arXiv preprint arXiv:2303.08774.
52. Van Eetvelde H, Mendonça LD, Ley C, et al. Machine learning methods in sport injury prediction and prevention: a systematic review. *J Exp Orthop*. 2021;8:1–15. https://doi.org/10.1186/s40634-021-00333-6.
53. Claudino JG, Capanema DO, de Souza TV, et al. Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: a systematic review, Sports Med. *Open*. 2019;5:1–12. https://doi.org/10.1186/s40798-019-0202-3.
54. Dean MC, Oeding JF, Diniz P, et al. Leveraging digital twins for improved orthopaedic evaluation and treatment. *J Exp Orthop*. 2024;11, e70084. https://doi.org/10.1186/s40634-024-00700-8.
55. Rodríguez P, Bautista MA, Gonzalez J, Escalera S. Beyond one-hot encoding: lower dimensional target embedding. *Image Vis Comput*. 2018;75:21–31.
56. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res*. 2011;12:2825–2830.
57. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 770–778.
58. Hou C, Lin X, Huang H, et al. Fossil image identification using deep learning ensembles of data augmented multiviews. *Methods Ecol Evol*. 2023;14:3020–3034.
59. Chorowski J, Wang J, Zurada JM. Review and performance comparison of svm-and elm-based classifiers. *Neurocomputing*. 2014;128:507–516.
60. Kotsiantis SB. Decision trees: a recent overview. *Artif Intell Rev*. 2013;39:261–283.
61. Sherstinsky A. Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. *Phys Nonlinear Phenom*. 2020;404, 132306.
62. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436–444.
63. Mahesh B. Machine learning algorithms-a review. *Int J Sci Res*. 2020;9:381–386 [Internet].
64. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *Adv Neural Inf Process Syst*. 2017;30.
65. Bengio Y, Ducharme R, Vincent P. A neural probabilistic language model. *Adv Neural Inf Process Syst*. 2000;13.
66. Soofi AA, Awan A. Classification techniques in machine learning: applications and issues. *J Basic Appl Sci*. 2017;13:459–465.
67. Maulud D, Abdulazeez AM. A review on linear regression comprehensive in machine learning. *J Appl Sci Technol Trends*. 2020;1:140–147.
68. Zhu Y, Wang Q, Wu C, et al. Logistic regression analysis on relationships between traditional Chinese medicine constitutional types and overweight or obesity. *Zhong xi yi jie he xue bao= Journal of Chinese Integrative Medicine*. 2010;8:1023–1028.
69. Yang C-C, Yen S-J, Chiu X-D, et al. Decision tree-based body constitution diagnosis system for traditional Chinese medicine. *Evid base Compl Alternative Med*. 2022; 2022, 5560087.
70. Chen X, Ma L, Chu N, Hu Y. Diagnosis based on decision tree and discrimination analysis for chronic hepatitis b in tcm. In: *2011 IEEE International Conference on Bioinformatics and Biomedicine Workshops (BIBMW)*. IEEE; 2011:817–822.
71. Li F, Zhao C, Xia Z, Wang Y, Zhou X, Li G-Z. Computer-assisted lip diagnosis on traditional Chinese medicine using multiclass support vector machines. *BMC Compl Alternative Med*. 2012;12:1–13.
72. Li G-Z, He Z, Shao F-F, Ou A-H, Lin X-Z. Patient classification of hypertension in traditional Chinese medicine using multi-label learning techniques. *BMC Med Genom*. 2015;8:1–6.
73. Zhang G, Huang Y, Zhong L, Ou S, Zhang Y, Li Z. An ensemble learning based framework for traditional Chinese medicine data analysis with icd-10 labels. *Sci World J*. 2015;2015, 507925.
74. Yuan F, Liu H, Chen S-q, Xu L. A clustering method of Chinese medicine prescriptions based on modified firefly algorithm. *Chin J Integr Med*. 2016;22: 941–946.
75. Xie D, Pei W, Zhu W, Li X. Traditional Chinese medicine prescription mining based on abstract text. In: *2017 IEEE 19th International Conference on E-Health Networking, Applications and Services (Healthcom)*. IEEE; 2017:1–5.
76. Tao Y, Chen YJ, Fu X, Jiang B, Zhang Y. Evolutionary ensemble learning algorithm to modeling of warfarin dose prediction for Chinese. *IEEE J Biomed Health Inf*. 2018;23:395–406.
77. Chen X, Ruan C, Zhang Y, Chen H. Heterogeneous information network based clustering for precision traditional Chinese medicine. *BMC Med Inf Decis Making*. 2019;19:1–12.
78. Yan J, Cai X, Chen S, Guo R, Yan H, Wang Y. Ensemble learning-based pulse signal recognition: classification model development study. *JMIR Med Inf*. 2021;9, e28039.
79. Huang C, Chen Y, Li B, et al. Traditional Chinese medicine constitution discrimination model based on metabolomics and random forest decision tree algorithm. *Evid base Compl Alternative Med*. 2022;2022, 3490130.
80. Wu Y, Pei C, Ruan C, Wang R, Yang Y, Zhang Y. Bayesian networks and chained classifiers based on svm for traditional Chinese medical prescription generation. *World Wide Web*. 2022:1–22.
81. Tian S, Wang J, Li Y, Xu X, Hou T. Drug-likeness analysis of traditional Chinese medicines: prediction of drug-likeness using machine learning approaches. *Mol Pharm*. 2012;9:2875–2886.
82. Zhang H, Li Z, Chen T, Liu J. Discrimination of traditional herbal medicines based on terahertz spectroscopy. *Optik*. 2017;138:95–102.
83. Sun X, Liu L, Wang Z, et al. An optimized multi-classifiers ensemble learning for identification of ginsengs based on electronic nose. *Sensor Actuator Phys*. 2017;266: 135–144.
84. Ming J, Chen L, Cao Y, Yu C, Huang B-S, Chen K-L. Rapid identification of nine easily confused mineral traditional Chinese medicines using Raman spectroscopy based on support vector machine. *J Spectr*. 2019;2019, 6967984.
85. Yang S, Shen Y, Lu W, et al. Evaluation and identification of the neuroprotective compounds of xiaoxuming decoction by machine learning: a novel mode to explore the combination rules in traditional Chinese medicine prescription. *BioMed Res Int*. 2019;2019, 6847685.
86. Han N, Qiao S, Yuan G, Huang P, Liu D, Yue K. A novel Chinese herbal medicine clustering algorithm via artificial bee colony optimization. *Artif Intell Med*. 2019; 101, 101760.
87. Tan L, Gu S, Wu C, Huang J, Zhou Y. K-means clustering method based on node similarity in traditional Chinese medicine efficacy. In: *2020 39th Chinese Control Conference (CCC)*. IEEE; 2020:742–747.
88. Li J-T, Wei Y-W, Wang M-Y, Yan C-X, Ren X, Fu X-J. Antibacterial activity prediction model of traditional Chinese medicine based on combined data-driven approach and machine learning algorithm: constructed and validated. *Front Microbiol*. 2021;12, 763498.
89. Tan CS, Leow SY, Ying C, et al. Comparison of ftir spectrum with chemometric and machine learning classifying analysis for differentiating guan-mutong a nephrotoxic and carcinogenic traditional Chinese medicine with chuan-mutong. *Microchem J*. 2021;163, 105835.
90. Li S, Yu Y, Bian X, et al. Prediction of oral hepatotoxic dose of natural products derived from traditional Chinese medicines based on svm classifier and pbpk modeling. *Arch Toxicol*. 2021;95:1683–1701.
91. Chen Z, Zhao M, You L, et al. Developing an artificial intelligence method for screening hepatotoxic compounds in traditional Chinese medicine and western medicine combination. *Chin Med*. 2022;17:58.
92. Sun J, Ni Q, Jiang F, et al. Discovery and validation of traditional Chinese and western medicine combination antirheumatoid arthritis drugs based on machine learning (random forest model). *BioMed Res Int*. 2023;2023, 6086388.
93. Wang Y, Ma L, Liao X, Liu P. Decision tree method to extract syndrome differentiation rules of posthepatitic cirrhosis in traditional Chinese medicine. In: *2008 IEEE International Symposium on IT in Medicine and Education*. IEEE; 2008: 744–748.
94. Zhou X, Li S, Tian F, et al. Building a disease risk model of osteoporosis based on traditional Chinese medicine symptoms and western medicine risk factors. *Stat Med*. 2012;31:643–652.
95. Li Y, Zhou H, Xie J, et al. A novel method for evaluating the cardiotoxicity of traditional Chinese medicine compatibility by using support vector machine model combined with metabonomics. *Evid base Compl Alternative Med*. 2016;2016, 6012761.
96. Kim S-K, Nam S, Kim S. Disarticle: a web server for svm-based discrimination of articles on traditional medicine. *BMC Compl Alternative Med*. 2017;17:1–6.

97. Yang B, Wang Y, Shan L, et al. A novel and practical chromatographic "fingerprint-roc-svm" strategy applied to quality analysis of traditional Chinese medicine injections: using kudiezi injection as a case study. *Molecules*. 2017;22:1237.
98. Yu L, Jin W, Li X, Zhang Y. Optimization of bioactive ingredient extraction from Chinese herbal medicine glycyrrhiza glabra: a comparative study of three optimization models. *Evid base Compl Alternative Med*. 2018;2018, 6391414.
99. Zhu J, Liu Y, Zhang Y, et al. Ihpreten: a novel supervised learning framework with attribute regularization for prediction of incompatible herb pair in traditional Chinese medicine. *Neurocomputing*. 2019;338:207–221.
100. Wang Y, Jafari M, Tang Y, Tang J. Predicting meridian in Chinese traditional medicine using machine learning approaches. *PLoS Comput Biol*. 2019;15, e1007249.
101. Tang Y, Li Z, Yang D, et al. Research of insomnia on traditional Chinese medicine diagnosis and treatment based on machine learning. *Chin Med*. 2021;16:1–21.
102. Lu W-W, Chen X, Ni J-L, et al. Study on the medication rule of traditional Chinese medicine in the treatment of acute pancreatitis based on machine learning technology. *Ann Palliat Med*. 2021;10, 106160625.
103. Tang ACY, Chung JWY, Wong TKS. Validation of a novel traditional Chinese medicine pulse diagnostic model using an artificial neural network. *Evid base Compl Alternative Med*. 2012;2012, 685094.
104. Zhang L, Zhang B. Non-invasive multi-disease classification via facial image analysis using a convolutional neural network. In: *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*. IEEE; 2019: 1383–1387.
105. Yao Y, Wang Z, Li L, et al. An ontology-based artificial intelligence model for medicine side-effect prediction: taking traditional Chinese medicine as an example. *Comput Math Methods Med*. 2019;2019, 8617503.
106. Song Z, Xie Y, Huang W, Wang H. Classification of traditional Chinese medicine cases based on character-level bert and deep learning. In: *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)*. IEEE; 2019:1383–1387.
107. Liu Z, Zheng Z, Guo X, et al. Attentiveherb: a novel method for traditional medicine prescription generation. *IEEE Access*. 2019;7:139069–139085.
108. Karki S, Ali AE, Alsadoon OH, Rashid TA. A novel solution of an enhanced error and loss function using deep learning for hypertension classification in traditional medicine. In: *2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA)*. IEEE; 2020:1–10.
109. Zhang Y, Wang L, Song J, Ji Z, Wu Q. Tcm medical record analysis algorithm based on recurrent convolutional neural network. In: *2021 36th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. IEEE; 2021:424–429.
110. Lan G, Li Y, Hu M, Sun Y, Zhang Y. Knowledge graph integrated graph neural networks for Chinese medical text classification. In: *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE; 2021:682–687.
111. Liang Z, Liu J, Ou A, Zhang H, Li Z, Huang JX. Deep generative learning for automated ehr diagnosis of traditional Chinese medicine. *Comput Methods Progr Biomed*. 2019;174:17–23.
112. Shi Q-Y, Tan L-Z, Seng LL, Wang H-J. Intelligent prescription-generating models of traditional Chinese medicine based on deep learning. *World J Trad Chin Med*. 2021; 7:361–369.
113. Wang X, Wang Y, Liu D, Wang Y, Wang Z. Automated recognition of epilepsy from eeg signals using a combining space-time algorithm of cnn-lstm. *Sci Rep*. 2023;13, 14876.
114. Patel Y, Shah T, Dhar MK, et al. Integrated image and location analysis for wound classification: a deep learning approach. *Sci Rep*. 2024;14:7043.
115. Lu X, Hu H, Li W, et al. Exploring hepatic fibrosis screening via deep learning analysis of tongue images. *J Tradit Complement Med*. 2024;14:544–549.
116. Cai C, Liu S, Wang L, et al. Classification of Chinese herbal medicine using combination of broad learning system and convolutional neural network. In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*. IEEE; 2019: 3907–3912.
117. Akter R, Hosen MI. Cnn-based leaf image classification for bangladeshi medicinal plant recognition. In: *2020 Emerging Technology in Computing, Communication and Electronics (ETCCE)*. IEEE; 2020:1–6.
118. Pudaruth S, Mahomoodally MF, Kissoon N, Chady F. Medicplant: a mobile application for the recognition of medicinal plants from the republic of Mauritius using deep learning in real-time. *IAES Int J Artif Intell*. 2021;10:938.
119. Jin Y, Ji W, Zhang W, He X, Wang X, Wang X. A kg-enhanced multi-graph neural network for attentive herb recommendation. *IEEE ACM Trans Comput Biol Bioinf*. 2021;19:2560–2571.
120. Bhuiyan MR, Abdullahil-Oaphy M, Khanam RS, Islam MS. Medinet: a deep learning approach to recognize bangladeshi ordinary medicinal plants using cnn. In: *Soft Computing Techniques and Applications: Proceeding of the International Conference on Computing and Communication*. IC3 2020. Springer; 2021:371–380.
121. Liu C, Xu F, Zuo Z, Wang Y. An identification method of herbal medicines superior to traditional spectroscopy: two-dimensional correlation spectral images combined with deep learning. *Vib Spectrosc*. 2022;120, 103380.
122. Gao W, Cheng N, Xin G, Khantong S, Ding C. Tcm2vec: a detached feature extraction deep learning approach of traditional Chinese medicine for formula efficacy prediction. *Multimed Tool Appl*. 2023;82:26987–27004.
123. Balasundaram A, Dilip G, Ashokkumar S, Manickam M, Gurunathan K, Kothandaraman D. Detecting true medicinal leaves among similar leaves using computer vision and cnn. *Natl Acad Sci Lett*. 2023;46:257–262.
124. Miao J, Huang Y, Wang Z, Wu Z, Lv J. Image recognition of traditional Chinese medicine based on deep learning. *Front Bioeng Biotechnol*. 2023;11, 1199803.
125. Quoc TN, Hoang VT. Traditional Vietnamese herbal medicine image recognition by cnn. In: *2023 15th International Conference on Knowledge and Smart Technology (KST)*. IEEE; 2023:1–5.
126. Deng N, Fu H, Chen X. Named entity recognition of traditional Chinese medicine patents based on bilstm-crf. *Wireless Commun Mobile Comput*. 2021;2021, 6696205.
127. Bai T, Guan H, Wang S, Wang Y, Huang L. Traditional Chinese medicine entity relation extraction based on cnn with segment attention. *Neural Comput Appl*. 2022: 1–10.
128. Yin Y, Zhang L, Wang Y, Wang M, Zhang Q, Li G-z. Question answering system based on knowledge graph in traditional Chinese medicine diagnosis and treatment of viral hepatitis b. *BioMed Res Int*. 2022;2022, 7139904.
129. Li X, Luo C, Chen K, et al. *Tlda: A Transfer Learning Based Dual-Augmentation Strategy for Traditional Chinese Medicine Rare Disease Syndrome Differentiation*. 2023.
130. Qi G, Zhang A, Xu Z, et al. Application of a complex network modeling approach to explore the material basis and mechanisms of traditional Chinese medicine: a case study of xuefu zhuyu decoction for the treatment of two types of angina pectoris. *IEEE Access*. 2022;10:114103–114117.
131. Chen H, Qin D, Zhang X, Zhang H, Liang X. A study on the classification of Chinese medicine records using bert, chest impediment as an example. In: *CCF International Conference on Natural Language Processing and Chinese Computing*. Springer; 2023: 29–37.
132. Xu C, Yuan F, Chen S. Research on assistant diagnostic method of tcm based on bert. In: *2021 11th International Conference on Information Technology in Medicine and Education (ITME)*. IEEE; 2021:282–286.
133. Liu Z, Luo C, Fu D, et al. A novel transfer learning model for traditional herbal medicine prescription generation from unstructured resources and knowledge. *Artif Intell Med*. 2022;124, 102232.
134. Liu Y, Yu J, Jia W, Liu Y. Tcm automatic diagnosis system based on knowledge graph and bert. In: *Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering*. 2023:829–834.
135. Chen Q, Ni J, Xu J, Gao X, Xia L. Generation of traditional Chinese medicine prescription driven by generative artificial intelligence gpt-4. *China Pharm*. 2023: 2825–2828.
136. Zhou X, Dong X, Li C, et al. *Tcm-ftp: Fine-tuning Large Language Models for Herbal Prescription Prediction*. 2024. arXiv preprint arXiv:2407.10510.
137. Wu H, Ji J, Tian H, et al. Chinese-named entity recognition from adverse drug event records: radical embedding-combined dynamic embedding-based bert in a bidirectional long short-term conditional random field (bi-lstm-crf) model. *JMIR Med Inf*. 2021;9, e26407.
138. Dou Y, Huang Y, Zhao X, et al. Shennongmgs: an llm-based Chinese medication guidance system. *ACM Trans Manag Inf Syst*. 2024.
139. Xue K, Zhou Y, Ma Z, Ruan T, Zhang H, He P. Fine-tuning bert for joint entity and relation extraction in Chinese medical text. In: *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE; 2019:892–897.
140. Wang Q, E H. A bert-based named entity recognition in Chinese electronic medical record. In: *Proceedings of the 2020 9th International Conference on Computing and Pattern Recognition*. 2020:13–17.
141. Song Y, Tian S, Yu L. A method for identifying local drug names in xinjiang based on bert-bilstm-crf. *Automat Control Comput Sci*. 2020;54:179–190.
142. Zhang M, Yang Z, Liu C, Fang L. Traditional Chinese medicine knowledge service based on semi-supervised bert-bilstm-crf model. In: *2020 International Conference on Service Science (ICSS)*. IEEE; 2020:64–69.
143. Guan Y, Li H, Xu W. A traditional Chinese medicine terminology recognition model based on deep learning: a tcm terminology recognition model. In: *Proceedings of the 6th International Conference on Big Data and Computing*. 2021:15–20.
144. Yang T-H, Pleva M, Hládek D, Su M-H. Bert-based Chinese medicine named entity recognition model applied to medication reminder dialogue system. In: *2022 13th International Symposium on Chinese Spoken Language Processing (ISCSLP)*. IEEE; 2022:374–378.
145. Zhang Y, Hao Y. Traditional Chinese medicine knowledge graph construction based on large language models. *Electronics*. 2024;13:1395.
146. Tong L, Tong X, Lei L, et al. Chinese text recognition and knowledge graph of shen nong ben cao jing based on bert pretrained language models. *Guidelines Stand Chin Med*. 2024;2:13–20.
147. Yao L, Jin Z, Mao C, Zhang Y, Luo Y. Traditional Chinese medicine clinical records classification with bert and domain specific corpora. *J Am Med Inf Assoc*. 2019;26: 1632–1636.
148. Hui Y, Du L, Lin S, Qu Y, Cao D. Extraction and classification of tcm medical records based on bert and bi-lstm with attention mechanism. In: *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE; 2020: 1626–1631.
149. Yang Z, Zhixiong Z, Huan L, Liangping D. Classification of Chinese medical literature with bert model. *Data Anal Knowledge Dis*. 2020;4:41–49.
150. Zhong X, Jia Y, Li D, Zhang X. Classification of acupuncture points based on the bert model. *J Data Anal Inf Process*. 2021;9:123–135.
151. Li Y, Peng X, Li J, et al. Relation extraction using large language models: a case study on acupuncture point locations. *J Am Med Inform Assoc*. 2024;31:2622–2631.
152. Zhou Z, Yang T, Hu K. Traditional Chinese medicine epidemic prevention and treatment question-answering model based on llms. In: *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE; 2023:4755–4760.
153. Zhu J, Gong Q, Zhou C, Luan H. Zhongjing: a locally deployed large language model for traditional Chinese medicine and corresponding evaluation methodology: a large language model for data fine-tuning in the field of traditional Chinese medicine, and a new evaluation method called tcmeval are proposed. In: *Proceedings of the 2023 4th International Symposium on Artificial Intelligence for Medicine Science*. 2023:1036–1042.

154. Yang G, Liu X, Shi J, Wang Z, Wang G. Tcm-gpt: efficient pre-training of large language models for domain adaptation in traditional Chinese medicine. *Comput Methods Programs in Biomed Update*. 2024, 100158.
155. Hua R, Dong X, Wei Y, et al. Lingdan: enhancing encoding of traditional Chinese medicine knowledge for clinical reasoning tasks with large language models. *J Am Med Inf Assoc*. 2024;31:2019–2029.
156. Tan Y, Zhang Z, Li M, et al. Medchatzh: a tuning llm for traditional Chinese medicine consultations. *Comput Biol Med*. 2024;172, 108290.
157. Zhang H, Wang X, Meng Z, Jia Y, Xu D. *Qibo: A Large Language Model for Traditional Chinese Medicine*. 2024. arXiv preprint arXiv:2403.16056.
158. Wang H, Liu C, Xi N, et al. *Huatuo: Tuning Llama Model with Chinese Medical Knowledge*. 2023. arXiv preprint arXiv:2304.06975.
159. Xiong H, Wang S, Zhu Y, et al. *Doctorglm: Fine-tuning Your Chinese Doctor Is Not a Herculean Task*. 2023. arXiv preprint arXiv:2304.01097.
160. Anil R, Dai AM, Firat O, et al. *Palm 2 Technical Report*. 2023. arXiv preprint arXiv:2305.10403.
161. Voigt P, Von dem Bussche A. In: *The Eu General Data Protection Regulation (Gdpr), A Practical Guide*. first ed. vol. 10. Cham: Springer International Publishing; 2017: 10–5555.
162. Felice F, Ley C, Groll A, et al. *Statistically Enhanced Learning: A Feature Engineering Framework to Boost (Any) Learning Algorithms*. 2023. arXiv preprint arXiv:2306.17006.
163. Cui C, Yang H, Wang Y, et al. Deep multimodal fusion of image and non-image data in disease diagnosis and prognosis: a review, Progress in Biomedical. *Engineering*. 2023;5, 022001.
164. Meskó B. The impact of multimodal large language models on health care's future. *J Med Internet Res*. 2023;25, e52865.
165. Liu F, Zhu T, Wu X, et al. A medical multimodal large language model for future pandemics. *NPJ Digital Med*. 2023;6:226.
166. Bhattacharya M, Pal S, Chatterjee S, Lee S-S, Chakraborty C. Large language model to multimodal large language model: a journey to shape the biological macromolecules to biological sciences and medicine. *Mol Ther Nucleic Acids*. 2024;35.
167. Tan X, Liu X, Shao H. Healthy China 2030: a vision for health care. *Value Health Reg Iss*. 2017;12:112–114.
168. Tan X, Zhang Y, Shao H. Healthy China 2030, a breakthrough for improving health. *Global Health Promot*. 2019;26:96–99.
169. Fielding JE, Kumanyika S, Manderscheid RW. A perspective on the development of the healthy people 2020 framework for improving us population health. *Public Health Rev*. 2013;35:1–24.
170. Raju TN. Achieving healthy people 2030 breastfeeding targets in the United States: challenges and opportunities. *J Perinatol*. 2023;43:74–80.
171. Kaur D, Uslu S, Rittichier KJ, Durresi A. Trustworthy artificial intelligence: a review. *ACM Comput Surv*. 2022;55:1–38.
172. Wu B, Li J, Yu J, et al. *A Survey of Trustworthy Graph Learning: Reliability, Explainability, and Privacy Protection*. 2022. arXiv preprint arXiv:2205.10014.
173. C. UNESCO. *Recommendation on the Ethics of Artificial Intelligence*. 2021.
174. N. N. G. A. I. G. S. Committee. *Ethical Norms for New Generation Artificial Intelligence Released*. Center for Security and Emerging Technology; 2021.
175. Schwabe D, Becker K, Seyferth M, Klaß A, Schaeffter T. The metric-framework for assessing data quality for trustworthy ai in medicine: a systematic review. *NPJ Digital Med*. 2024;7:203.
176. Chaddad A, Peng J, Xu J, Bouridane A. Survey of explainable ai techniques in healthcare. *Sensors*. 2023;23:634.
177. Holzinger A, Biemann C, Pattichis CS, Kell DB. *What Do We Need to Build Explainable Ai Systems for the Medical Domain?*. 2017. arXiv preprint arXiv:1712.09923.