International Research Journal of Multidisciplinary Scope (IRJMS), 2026; 7(1): 1063-1075 IRJMS

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2026.v07i01.08145

Too Early to Call? - A Critical Evaluation of AI's Nascent Role in Comprehension for Traditional Chinese Medicine

Chengxi Huang*

Culver Academies, 1300 Academy Rd, Culver, IN 46511, USA. *Corresponding Author's Email: chengxi.huang07@gmail.com

Abstract

Traditional Chinese Medicine (TCM) is a system of medical concepts, diagnostic methods, and therapies with a rich and ancient history in China. While TCM has long been a cornerstone of healthcare in Chinese communities, its influence and popularity are much less realized on a global scale. The public's understanding of TCM can be largely affected by the most common sources of accessible information online, especially with the emergence and quick development of AI modern technologies in recent years, as these tools are potentially becoming a primary source of knowledge for many, thus posing a significant impact on how people perceive and learn about TCM. This raises an important question about the accuracy of AI-generated responses and the potential consequences of misrepresenting TCM in public. This study evaluated the accuracy of AI responses on TCM research by comparing them to bibliometric analysis, which is traditionally used to analyze academic publications and trends. The findings show that AI is effective at identifying broad, general patterns in TCM research. However, it often lacks the nuanced details and specific insights that a thorough bibliometric analysis can provide. Despite these limitations, the study still revealed a promising synergy between the two approaches. AI can assist bibliometric analysis by helping quickly capture general research categories that would be time-consuming to identify manually. This suggests that integrating the power of AI with traditional bibliometric methods could lead to unexpected benefits, offering a more efficient and comprehensive way to understand the complex landscape of TCM research.

Keywords: AI Chatbots, Bibliometric Analysis, Research Theme, Traditional Chinese Medicine.

Introduction

Traditional Chinese Medicine (TCM) is a general term that refers to medical concepts, clinical diagnostic methods, and treatments that originated in ancient China thousands of years ago. The Yellow Emperor's Medical Classic (Huang Di Nei Jing) has been considered the oldest text, being compiled about two thousand years ago and laying the foundation of TCM (1). TCM plays a crucial role in Chinese medical treatment and is often combined with modern therapies. Its central concept is Qi, energy within all living beings, as well as yin and yang, two opposite and combating forces derived from Qi (2). TCM believes that all illnesses result from the imbalance between yin and yang, and a key practical concept often used is the network of acupoints, specific regions in the human body that are crucial to their balance (3).

Diagnosis in TCM generally relies on 'Four Examinations': observing, listening/smelling, asking, and feeling. Observing involves careful inspection of patients' physical features (4). Listening and smelling focus on sounds and odors from signs like patients' cough, sneeze, and body excrement, etc. The asking method is a detailed query about patients' symptoms, such as the location and intensity of pain. Lastly, feeling or palpation primarily refers to taking pulse, as it is considered a key indicator of internal imbalance that might provide clues to therapies (5). Based on the diagnostic results, TCM treatments are often a combination of different therapies, which can vary in effects and forms but are normally categorized as herbal medicine, acupuncture, moxibustion, massage or Tuina, and exercise like Tai Chi (6). Many TCM texts record different herbs and their corresponding effects on the human body, and herbal medicines are often used with other treatments, especially assisting recovery and regulating Qi. Acupuncture involves inserting needles into acupoints on the human body to regulate Qi and restore the balance of yin and yang. Moxibustion is often utilized to remove excess yin from the human body by placing burning mugwort leaves either directly on or above regions of the body (7). Tuina refers to stimulating internal responses in the human body and restoring

(Received 11th September 2025; Accepted 27th December 2025; Published 28th January 2026)

internal balance through massaging acupoints. Tai Chi is an exercise that incorporates physical movement, mental focus, and breath regulation to benefit the body and mind. Taken together, TCM offers a diverse range of treatments to restore internal balance and promote health, and a key benefit of this medical system is its non-invasive approach, as it does not rely on surgical procedures.

While TCM has been practiced in China for thousands of years, it has yet to gain worldwide recognition (8, 9). Public perception about TCM is normally from oral transmission, despite that much research on TCM has been conducted, which leads to mystery and pseudoscience (10). In contrast to TCM's prevalent use and popularity in Chinese communities, most other societies remain largely unfamiliar with it and may have misconceptions regarding TCM. Chinese researchers and medical professionals have put their efforts into increasing the global influence of TCM. However, the public often relies on search engines like Google for information, potentially leading to misconceptions about TCM. The emergence and increasing prevalence of AI tools in recent years could possibly exacerbate this issue further.

Previous studies have reviewed TCM research trends using bibliometric approaches, particularly those examining publication networks, topic evolution, and international collaboration patterns (11, 12). Recent works have also explored AI's potential role in applying machine learning to syndrome differentiation and modeling multi-herb interactions using network pharmacology (13, 14). However, very few studies have systematically compared AI-generated research trends with formal bibliometric outputs. It is imperative to acquire a deep understanding of the discrepancy brought by reliance on AI. This research addresses this gap by integrating topic modeling with AI-based thematic prediction, thereby situating the present study within two converging domains: TCM informatics and AI-assisted medical knowledge evaluation.

Further, to evaluate the growing reliance on AI as a primary source of public knowledge about TCM, this study establishes a direct comparison between AI-generated interpretations and traditional bibliometric evidence. Specifically, the study pursued four research objectives: characterizing long-term publication trends in TCM research with bibliometric methods; mapping international collaboration patterns and identifying dominant contributors; extracting major TCM research topics with LDA topic modeling; and assessing the accuracy with which a state-of-the-art AI model (ChatGPT) captures these same patterns. These objectives guided the structure of the entire study, ensuring that each analytical component corresponds directly to the overarching questions posed.

Methodology

Data Source

The data used in this study were obtained from the publicly accessible database of PubMed. A search for the term "Traditional Chinese Medicine" in the abstracts was performed to extract publication records of interest. The search results were manually checked by a random survey of at least 100 records from the top and bottom to validate their relevance. The search results with available abstracts were up to January 2025 and downloaded in the "PubMed" format for subsequent analyses.

Data Analysis

Bibliometric methods are widely used to map research productivity, collaboration structures, and thematic evolution across scientific fields (15, 16). Information from the downloaded publication records was extracted using Python to fit the three categories of bibliometric analysis focus as depicted in Figure 1. The number of TCM-related articles from 1950 to January 30th, 2025, was used to identify long-term trends in the volume of publications. To assess the global recognition of TCM research, the analysis examined the geographical distribution of published TCM research, and a collaborative network was created using the information of co-authors' institutional affiliations to visualize international collaboration in TCM research. Keywords from the abstracts were used to perform Latent Dirichlet Allocation (LDA) topic modeling to reveal underlying TCM research themes and focuses. This unsupervised machine learning tool has often been used for the identification of latent topics and their distributions in a large corpus (17).

Because PubMed abstracts are primarily in English, no Chinese segmentation tools were applied in the core LDA pipeline. However, to

account for the influence of traditional TCM terminology, a custom dictionary was compiled to include Pinyin transliterations (e.g., “qi”, “yin-yang”, “tuina”, “moxibustion”, “tai chi”). These terms were preserved during preprocessing to avoid fragmentation of traditional phrasing. No multilingual alignment models were used in the primary analysis, though future improvements could incorporate cross-lingual embeddings to better bridge classical Chinese terminology with English biomedical descriptors.

To qualitatively assess the accuracy of generated themes and the bibliometric results, a structured human evaluation procedure was conducted. A single trained annotator with knowledge in TCM and experience in literature-based classification reviewed a stratified sample of 120 abstracts drawn from the major topics identified through LDA. The annotator applied a predefined annotation framework consisting of four dimensions: disease domain, intervention type, mechanistic focus, and study type. Example annotation items included “herbal formula targeting inflammatory cytokines”, “acupuncture analgesia mechanism”, and “TCM-microbiota interactions”. Although multiple raters were not available in this study, internal reliability was evaluated through a test-retest calibration, in which the annotator revisited all the abstract samples for re-evaluation, achieving a self-consistency agreement score over 0.8. This procedure ensured consistent application of the annotation rules and enhanced transparency regarding the qualitative assessment process.

AI Queries

Prompts regarding the same aspects of TCM research as visualized in the above data analysis were provided to ChatGPT 4.0 to evaluate the AI generative accuracy (18). The AI responses were generated in July 2025. For the publication trend, accuracy would be based on the publication volumes in different periods and significant years suggested by AI. For research topics, it examined the number of overlaps between the topics determined by LDA topic modeling and by AI. For international collaboration, the top collaborators suggested by AI and by the bibliometric analysis were compared for similarities. These comparisons would reveal the complementary relations between AI and traditional research methods. This, in turn, would help identify potential areas to amend public misinformation about TCM and improve research methodologies in the future.

This diagram in Figure 1 illustrates the overall workflow, including: (i) data collection from PubMed; (ii) preprocessing of publication records; (iii) three analytical branches—publication trend analysis, international collaboration mapping, and LDA topic modeling; (iv) comparison of bibliometric results with AI-generated themes. Each step is represented with directional arrows to show analytical progression.

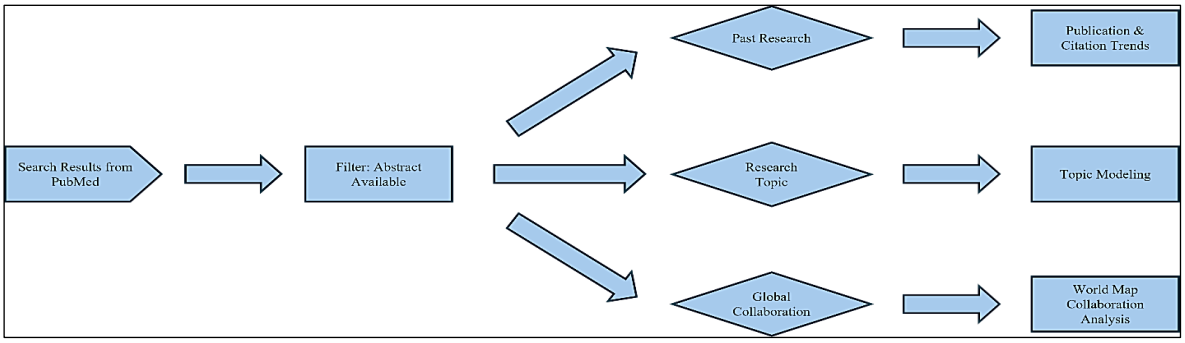


Figure 1: Workflow of the Study Design

Results

General Publication Trends

A total of 136,081 TCM publications were found in the timespan from 1950 to January 2025. The majority of these, 129,900, were co-authored, reflecting a high level of collaboration with an average index of 6.924. The publication trend, as shown in Figure 2, suggests a gradual development before 2010, with average numbers of publications of less than 2,000, and a sharp increase to more than 10,000 articles after 2010. The number of publications increased by about 4,000 articles from 2021 to 2022, reaching 21,560. However, it only increased by less than a hundred in 2023, but reached a new peak again with 28,170 articles in

2024, increasing by almost 7,000 articles. Overall, publications have increased significantly in recent decades, suggesting a growing interest and dedication to research in TCM.

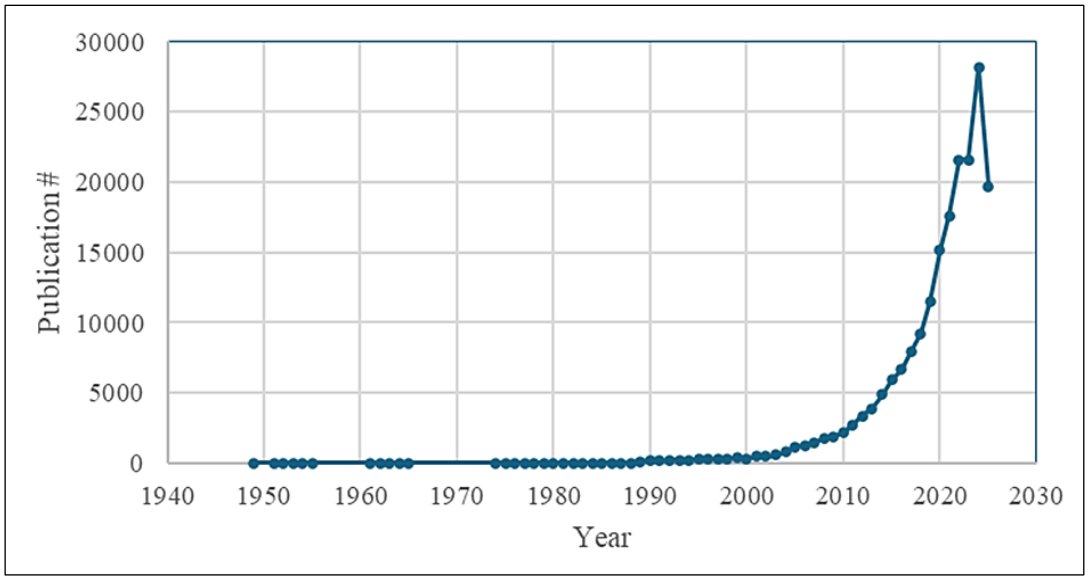
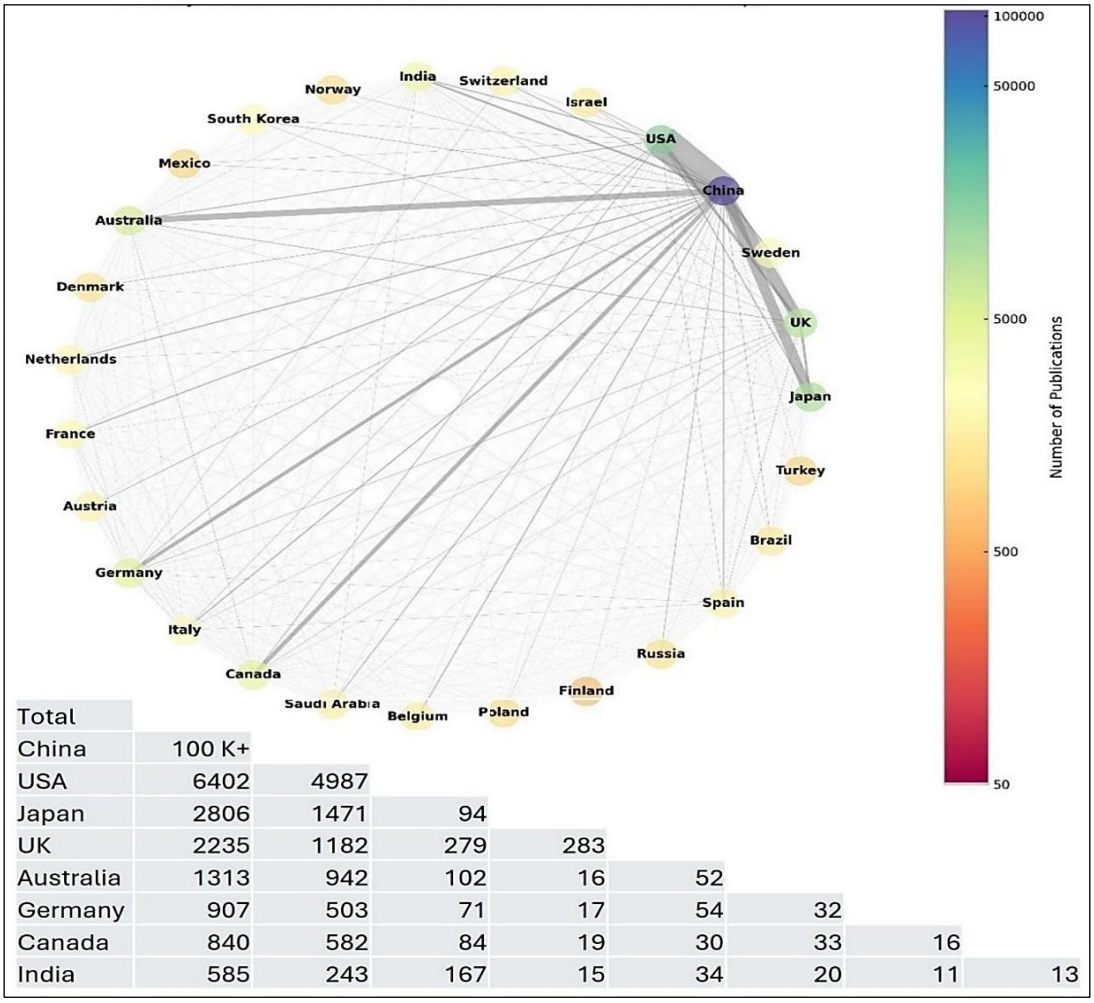


Figure 2: Number of Publications per Year from 1950 to 2025 (January)



Total								
China	100 K+							
USA	6402	4987						
Japan	2806	1471	94					
UK	2235	1182	279	283				
Australia	1313	942	102	16	52			
Germany	907	503	71	17	54	32		
Canada	840	582	84	19	30	33	16	
India	585	243	167	15	34	20	11	13

Figure 3: International Collaboration Network (The Color of Each Nation Represents Its Total Publications, and the Thickness of Each Line Represents the Number of Collaborations)

International Collaboration

Among the 17 countries identified in the publications associated with TCM, China ranks first with more than 100,000 publications, as shown in Figure 3. The USA ranks second with 6,402 publications, and Japan follows with 2,806 publications. China collaborated with the USA on 4,987 articles and with Japan on 1,471 articles. Notably, although Japan has a higher number of total publications, it has fewer collaborations with all other top nations except China. Overall, TCM research involved many different nations located on different continents, yet collaborations were heavily centered around China.

LDA Topic Modeling

To determine the trends in TCM research, unigram and bigram LDA models were both created and trained. Generic terms, common English terms, and search terms were used as stop words to clarify the trends demonstrated in the models. The optimum number of topics in each model was determined by the highest coherence score among models, ranging from 2 to 20 topics. The optimal numbers of topics were 5 and 4 for the unigram and bigram models, respectively, as shown in Figure 4. The two models showed independent and overlapping topics regarding TCM, as demonstrated in Figure 5 and Figure 6.

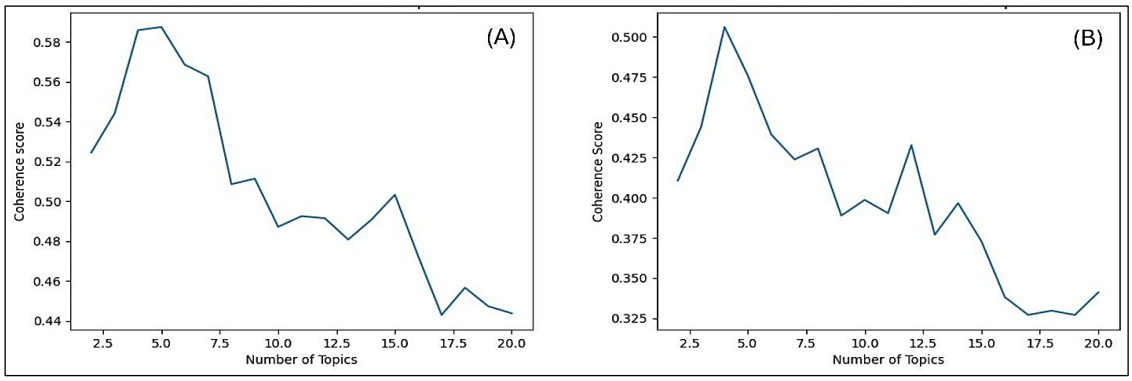


Figure 4: Model Coherence Scores and Number of Topics (A) Unigram Model; (B) Bigram Model

Unigram Model Results: The unigram model identified five topics and their frequent terms as shown in Figure 5. The H indices were shown to validate the popularity and impact of the identified topics. This index is a value computed through the number of publications and citations to suggest the impact of a researcher or topic (19). Topic 5 had the highest H index. It suggests that the articles related to topic 5 were cited more than those related to the other identified topics, although topic 5 did not appear in the abstracts as frequently as the other identified topics did. A graph with the topic trends was created and indicated a significant increase in publication numbers in 2000, 2010, and 2020, as shown in Figure 5. Some small peaks were also visible in the number of publications around 1980 and 1990.

Bigram Model Results: Four topics were identified in the bigram model, as listed in Figure 6. The bigram model identified similar topics to the unigram model, but the topic sizes and orders were opposite for "Acupuncture Treatment and Pain" and "TCM Treating Tumor and Cancer Through Gene Expression." Again, all of the topics demonstrated a significant increase in the number of publications after 2000, 2010, and 2020. Interestingly, topic 1 in this model experienced a sharp increase in the number of publications between 1980 and 1990.

Inter-topic distance maps generated during the topic modeling were used to examine the sizes and distances between different identified topics, which represent the overall prevalence and relatedness of topics (20). The identified topics in both models are relatively distant from each other and demonstrate no overlap, as demonstrated in Figure 7.

Topic Composite: The topics from the unigram and bigram models are merged to show the overall major research themes in TCM. These topics were examined for their relative impact through the Web of Science citation report analysis, as depicted in Table 1. Although topic sizes correspond to the publication volume, H-indices suggest that some topics are less prevalent but more focused and recognized.

#	Topic	Frequent Terms
1	TCM treating tumor and cancer through gene expression	Cancer, cells, expression, genes, pathway, tumor, targets, signaling, proliferation
2	TCM treating inflammation	Induced, inflammatory, expression, cells, beta, immune, inflammation, injury, protein
3	TCM treating diabetes/hypertension and effects on gut microbiota	Diabetes, treatment, blood, liver, serum, gut, hypertension, microbiota, alpha
4	TCM and metabolites and herb	Compounds, components, herb, acid, active, drug, activity, metabolites, chemical
5	Acupuncture treatment and pain	Treatment, risk, acupuncture, included, efficacy, factors, rate, score, pain

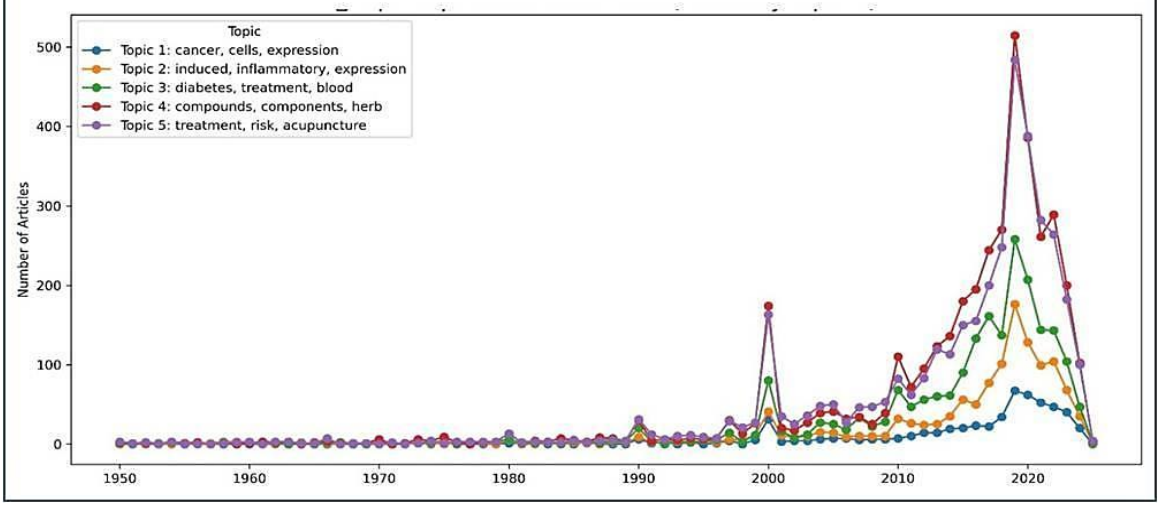


Figure 5: Publication of Topics from the Unigram Model

#	Topic	Frequent Terms
1	Acupuncture treatment and pain	treatment, acupuncture, pain, stroke, risk, higher, depression, model, infection, used
2	TCM treating inflammation	inflammation, treatment, liver, expression, induced, increased, model, reduced, decreased, fibrosis
3	TCM and metabolites and herb and obesity	compounds, obesity, metabolites, used, components, drug, new, application, products, properties
4	TCM treating tumor and cancer through gene expression	cancer, tumor, expression, cells, treatment, genes, targets, mir, infection

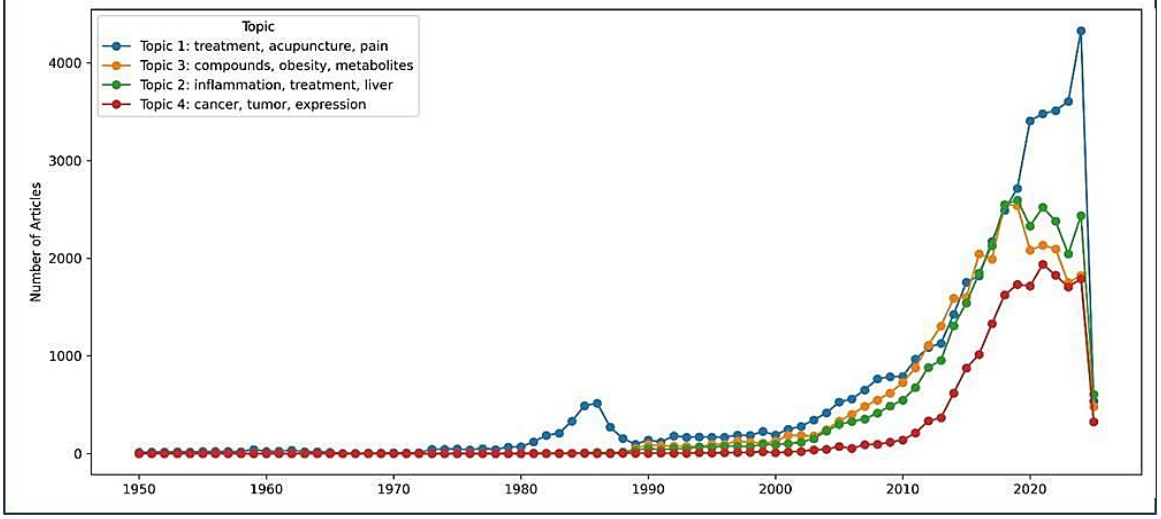


Figure 6: Publication of Topics from the Bigram Model

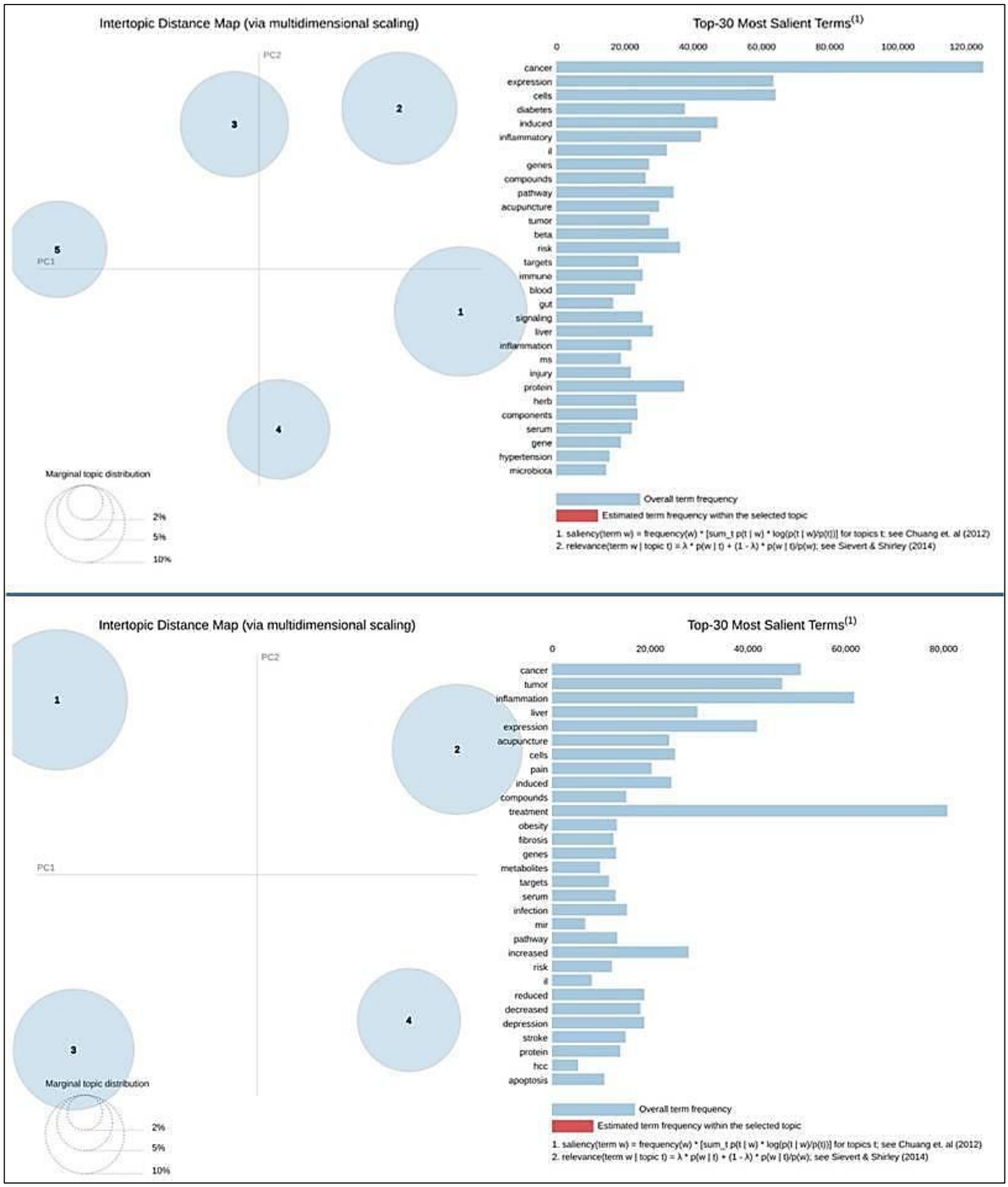


Figure 7: Inter-Topic Distance Map for the Unigram (Upper) and Bigram (Lower) Models

Table 1: Overall Topics and their Impact

#	Topics	H Index
1	Acupuncture treatment and pain	148
2	TCM treating inflammation	47
3	TCM treating obesity	13
4	TCM treating tumor and cancer through gene expression	61
5	TCM treating diabetes	30
6	TCM treating hypertension	65
7	TCM and Microbiota	73
8	TCM herbs and metabolites	60

ChatGPT Comparisons

Publication History:

Question: What is the research history of Traditional Chinese Medicine? Has it increased or decreased, and by how much, in different periods? Include significant periods and potential reasons.

Response:

General Trend: TCM research grew steadily from the 1950s, expanded rapidly after 2000, and peaked globally between 2015 and 2025.

Key Points and Reasons:

a) 1950s–1990s: Slow growth; TCM was supported by the Chinese government and integrated into national health policy, but research remained largely domestic and descriptive due to limited scientific tools and international isolation.

b) 2000–2010: Rapid increase in publications (from ~500 to 2,000/year); driven by global interest in alternative medicine, improved research infrastructure, and adoption of clinical trials and pharmacology.

c) 2010–2019: Output exceeded 5,000 publications/year; AI, network pharmacology, and genomics enhanced research depth. Government funding (e.g. 973 Program) and global collaboration grew significantly.

d) 2020–2025: Peak period (10,000+ publications/year); COVID-19 spotlighted TCM in China's pandemic response, leading to a surge in respiratory and immune-related studies. Push for international standards and digital innovation accelerated.

International Collaboration:

Question: What countries have participated the most in TCM research and collaborated with other nations?

Response:

From 1950 to 2025, China has been the clear global leader in Traditional Chinese Medicine (TCM) research, contributing the majority of publications and initiating most international collaborations. Its dominance grew significantly after 2000, driven by strong government support, integration of TCM in national health policy, and institutional expansion. The United States has been China's top research partner, particularly in areas like acupuncture, herbal medicine, and integrative health, often supported by NIH and academic partnerships. Germany, the UK, Australia, and Canada have also played major roles in evidence-based TCM studies, focusing on clinical trials, safety evaluations, and regulatory integration. In Asia, Japan and South Korea have contributed extensively to TCM and related systems (like Kampo and Korean medicine), with growing co-authorship ties to China. More recently, countries like India, Switzerland, Sweden, and Brazil have increased participation, especially in pharmacological and cross-cultural medical research. Over time, collaboration has shifted from limited bilateral exchanges to a dense global research network centered around China.

Research Topics:

Question: What are some hotspots of TCM research, such as different TCM concepts or TCM in treating different diseases?

Response:

TCM Concepts and Theories

a) Syndrome Differentiation (*bianzheng lunzhi*): Central to diagnosis and treatment; studies aim to link syndromes with biomarkers or disease subtypes.

b) Qi, Yin-Yang, and Five Elements: Research attempts to interpret these traditional ideas through systems biology, neuroendocrine pathways, and energy metabolism.

c) Herbal Compatibility and Formula Theory: AI and network pharmacology are used to analyze multi-herb interactions and their pharmaco-dynamics.

Disease Treatment Hotspots

a) Respiratory Diseases

- Especially active since COVID-19 (2020–2025).
- Formulas like Lianhua Qingwen and Jinhua Qinggan studied for antiviral, anti-inflammatory, and immune-modulating effects.

b) Cancer

- Use of TCM as adjuvant therapy alongside chemotherapy/radiation.
- Research focuses on herbal formulas for reducing side effects, improving immunity, and slowing tumor progression (e.g., YIV-906 in clinical trials).

c) Chronic Pain and Musculoskeletal Disorders

- Acupuncture widely researched and integrated into guidelines for conditions like low back pain, osteoarthritis, and migraines.
- Mechanisms studied through fMRI and neurochemical modulation.

d) Digestive Disorders
- High volume of research on functional dyspepsia, IBS, ulcerative colitis.
- Herbal treatments often target gut microbiota, inflammation, and motility.

e) Neurological Disorders
- Studies on stroke recovery, Alzheimer's, Parkinson's, and insomnia.
- TCM therapies investigated for neuroprotection and cognitive enhancement.

f) Metabolic Diseases
- Focus on type 2 diabetes, obesity, and metabolic syndrome.
- Herbs/formulas like Berberine, Jinqi Jiangtang used for glycemic control and insulin sensitivity.

ChatGPT-Bibliometric Comparison:

Regarding publication history, ChatGPT's response aligns with the bibliometric analysis on the general trends in each period. For instance, the 1950s to 1990s were marked by "slow growth," and 2020-2025 was the "peak period." However, AI could not identify specific significant points, such as the flat from 2022 to 2023. Overall, ChatGPT suggested a general trend very similar to the results from the bibliometric analysis.

In terms of international collaboration, ChatGPT and the bibliometric analysis partially overlap in this aspect of TCM research. First, both of them identified China as the dominant nationality in TCM research and the United States as its top collaborator. However, for other significant countries such as Japan, the UK, and Canada, ChatGPT only stated that they either "played major roles" or "have contributed extensively to TCM," but did not specify differences between these nations. For example, Canada and Australia have similar numbers of total publications, but Canada has fewer international collaborations, as shown in Figure 3. Notably, ChatGPT suggested that China's dominance in publications can be attributed to "strong government support" and "integration of TCM in national health policy." In addition, for Germany, the UK, Canada, and Australia, ChatGPT indicated that they "played major roles in evidence-based TCM studies," which use evidence-based medicine (EBM) to standardize the evaluation of TCM and have largely transformed TCM research in the past 20 years (21). Such background information was not seen through bibliometric analysis.

For research topics, ChatGPT identified three concepts and six categories of diseases as the hotspots of TCM research. The ChatGPT response, to different degrees, reflects seven of the eight topics captured by the bibliometric analysis. AI suggested that three hotspots of TCM concepts are "Syndrome Differentiation," "Qi, Yin-Yang, and Five Elements," and "Herbal Compatibility and Formula Theory." Examining the topics from the bibliometric analysis, syndrome differentiation, and concepts like Qi and Yin-Yang do not appear to be keywords or hotspots. However, the third concept of herbal medicine does overlap with topics 3 and 8 as depicted in Table 1. Among the six disease categories mentioned by ChatGPT, "Cancer" and "Chronic Pain and Musculoskeletal Disorders" are the two diseases also seen as independent topics (topic 4 and topic 1, respectively) in the bibliometric analysis. "Respiratory Diseases," "Digestive Disorders," and "Metabolic Diseases" are three disease categories that are not directly identified as topics in the bibliometric analysis but include descriptions that overlap with identified topics. Under "Respiratory Diseases," "anti-inflammatory" and "immune-modulating effects" are related to topic 2, "TCM Treating Inflammation." Under "Digestive Disorders," "gut microbiota" and "inflammation" are keywords related to topic 2 and topic 7 (TCM and Microbiota). Under "Metabolic Diseases," keywords include "diabetes," "obesity," and "herb," corresponding to topic 5 (TCM Treating Diabetes), topic 3 (TCM and Metabolites and Herbs and Obesity), and topic 8 (TCM and Metabolites and Herbs). The only topic not mentioned in the ChatGPT response is topic 6, "TCM Treating Hypertension." Overall, ChatGPT reflects a surprisingly high portion of the topics identified by the bibliometric analysis.

Discussion

The purpose of this study is to examine TCM research publication trends and hotspots while evaluating AI's ability to identify these patterns. To assess AI's accuracy, a bibliometric analysis was conducted using PubMed as the database and focused on three aspects of TCM: publication trends, international collaboration, and research topics. ChatGPT was given questions regarding the same aspects of TCM, and its response was

compared with the results from the bibliometric analysis.

The bibliometric analysis shows that the TCM research publication has experienced significant growth since the 21st century, which can be attributed to the fact that TCM was largely limited to China before the internet became a convenient tool for researchers. In recent years, the sharp increase in the number of annual publications is likely a result of the COVID-19 pandemic, in which TCM was employed during the pandemic and in the recovery period post-pandemic (22). This implies that the recent sharp increase in TCM research may be temporary and that the rate of increase may fall back down unless significant new efforts are dedicated to TCM research in the future.

According to the bibliometric analysis, the international collaboration of TCM research heavily centers on China. Other top contributors collaborating with China include the U.S., Japan, the U.K., Australia, Germany, and Canada. Collaborations between nations without China remain comparatively scarce. These results indicate that TCM research and collaborations were largely driven by Chinese researchers. More independent publications and collaborations without China's participation are necessary towards a global acceptance of TCM.

LDA modeling results are examined along with the H index from the Web of Science. Topic 1 stands out with the highest H index and reaches multiple peaks. Around the 1970s and 1980s, treating musculoskeletal disorders like myofascial pain syndrome revealed the potential of Chinese acupuncture and led to increasing research about it (23). The continuous popularity of acupuncture worldwide can be attributed to its efficacy in pain relief and its wide range of applications (24). Topic 2 has a much lower H index and experiences more rapid growth in years following the COVID-19 pandemic, which is likely attributed to the increase in popularity of the anti-inflammation effects of TCM in recent research. Topics 3 and 8 are very similar, and this study examines them separately to determine whether adding "obesity" as a keyword would have a significant effect on the H index. As a result, topic 8 does have a much higher H index than topic 3 does, suggesting that "obesity" might appear often in research but is cited much less often than "herbs" and "metabolites" in research. Topics 4, 5, 6, and 7 have relatively similar H indices and demonstrate similar trends as the overall trend of TCM research publication, and chronic diseases like diabetes and hypertension are categories where TCM has often been employed (25). Based on these results, the key factor that determines the popularity of different topics is likely an incident that can bring TCM into the public eye. Acupuncture has been the most popular TCM topic on a global scale because of the research in musculoskeletal disorders in the late 20th century, which brought acupuncture to nations outside of China much earlier than when other aspects of TCM were introduced. Chronic diseases have been studied for a longer history without TCM, explaining the comparatively lower popularity of TCM in treating them. To this end, TCM researchers can focus on new and trending medical topics that are not limited to China, inspiring interests in TCM in other nations and bringing attention to TCM.

After comparing ChatGPT's response and the results of the bibliometric analysis, it can be seen that AI has a mature understanding of the overall TCM publication trend. ChatGPT also successfully identified the major countries that participated in international collaborations and the hotspots of TCM research. However, its response lacks certain details when compared to the bibliometric analysis. For the overall publication trend, AI only gives out general trends for each period but fails to identify and explain significant twists within a period, such as the drop from 2022 to 2023. For the international collaboration aspect, AI accurately identified the top two nations and a list of active nations, but cannot differentiate those active nations in the same way as the bibliometric analysis does. The research topic aspect is particularly interesting because AI's response may seem inaccurate at first glance, but actually reflects almost all the topics identified in the bibliometric analysis. A potential factor that might have contributed to this phenomenon is that disease categories are not often written in published TCM research, which means that although certain topics might be closely connected to a type of disease, the bibliometric analysis might only reflect the specific disease instead of the general type. Therefore, examining AI's response and the bibliometric analysis together may reveal trends that are hard to identify using just one tool. For example, the bibliometric analysis suggests that anti-

inflammation has been a hotspot, while ChatGPT adds on to that by indicating anti-inflammation in respiratory diseases, offering researchers a possibility that the topic of anti-inflammation is a hotspot because of TCM research in respiratory diseases. Without ChatGPT, it might be difficult for researchers to identify respiratory diseases among all the types of diseases with inflammatory responses. Although AI on its own has some flaws when compared to the bibliometric analysis, incorporating AI in traditional bibliometric analysis may bring unexpected benefits and uncover details difficult to discover.

On the other hand, AI-generated responses could potentially contain three principal categories of biases: (i) misinterpretation – where AI overgeneralizes TCM explanatory frameworks, e.g., interpreting traditional concepts like "Qi" through systems biology, neuroendocrine pathways, or energy metabolism without textual evidence; (ii) hallucinated citations or unsubstantiated claims – where AI references programs (e.g., "973 program influence on TCM research output") or clinical trial outcomes that did not appear in the bibliometric dataset. These claims may reflect true background knowledge but are unsupported by the analyzed corpus; (iii) improper topic mappings – where AI assigns disease categories (e.g., respiratory disorders, digestive disorders) that do not explicitly appear in the LDA topics but are implicitly connected through shared mechanisms like inflammation or microbiota regulation. Such limitations align with documented hallucination and overgeneralization behaviors in large language models used for scientific and medical applications (26, 27). Again, these observations demonstrate that AI excels in high-level conceptual grouping but may introduce unsupported details when specifics are missing. Further, an important conceptual consideration is epistemic plurality, the coexistence of multiple valid frameworks for understanding health and disease. TCM operates under holistic, pattern-based reasoning rooted in classical doctrines, whereas AI systems trained predominantly on biomedical literature may impose reductive interpretations framed around biochemistry or molecular pathways. Ensuring that AI technologies preserve TCM's conceptual integrity, rather than force-mapping it into Western biomedical constructs, is essential for culturally respectful and accurate representation. Balancing predictive accuracy with interpretability and cultural validity remains a central challenge in medical AI applications (28, 29). Future AI models should incorporate classical TCM texts, bilingual corpora, and culturally grounded annotation frameworks to maintain epistemic balance.

In terms of limitations, the bibliometric portion of this study is based on only the PubMed database, as it is open to the public and more relevant to the medical field. Other popular sites for research publication, like the Web of Science and Scopus, might provide other perspectives on TCM research. ChatGPT was the only AI chatbot examined since it is also the most popular AI chatbot worldwide, but its responses can vary depending on the model and the prompts given. A key consideration arising from this dual methodology is the significant representation mismatch across the compared systems. Bibliometric analysis relies exclusively on English-language scientific abstracts that disproportionately emphasize mechanistic, biomedical, and evidence-based terminology. In contrast, ChatGPT draws on a far broader cross-domain corpus—including textbooks, popular science writing, multilingual content, and non-biomedical sources—which introduces domain-adaptation effects. This mismatch partly explains why AI prioritizes classical doctrinal concepts (e.g., Qi, syndrome differentiation) that are underrepresented in English publications but prominent in Chinese-language scholarship. Future work should incorporate domain-adapted models fine-tuned on bilingual TCM corpora to reconcile these representation gaps. Future work should address these limitations by incorporating additional authoritative databases to increase sample size, integrating other leading AI chatbots for comparative analysis, and, most crucially, employing domain-adapted AI models fine-tuned on bilingual TCM corpora to reconcile these observed representation gaps. Furthermore, in-depth content analysis of articles from the identified significant publication periods should be pursued to validate the potential reasons for observed trends.

Conclusion

This study systematically compared AI-generated interpretations of TCM research with patterns derived from bibliometric analysis. While ChatGPT successfully identified major publication trends, leading countries, and high-impact research themes, the AI model also introduced misinterpretations and unsupported generalizations due to representation mismatches and conceptual differences between TCM and biomedical literature. Bibliometric analysis remains essential for precise, data-driven characterization of research landscapes, but AI demonstrates clear value in generating broader contextual connections and highlighting conceptual domains not explicitly represented in abstracts. Together, the two approaches provide a complementary framework for understanding the evolving field of TCM research. Moving forward, the AI-supported TCM research comprehension should advance through an integrated roadmap focusing on accuracy and cultural alignment. Immediate efforts should prioritize the creation of domain-adapted AI models, improved bilingual terminology datasets, and standardized evaluation protocols. This foundation will enable medium-term objectives, including the implementation of human-in-the-loop systems, cross-model consistency testing, and the critical incorporation of Chinese-language and classical TCM corpora. Ultimately, the long-term goal is to establish regulatory guidelines, cross-cultural validation frameworks, and education systems that ensure AI tools operate ethically, accurately, and in full alignment with both biomedical evidence and traditional TCM epistemology.

Abbreviations

LDA: Latent Dirichlet Allocation, TCM: Traditional Chinese Medicine.

Acknowledgment

The author would like to sincerely thank editors and reviewers for their constructive suggestions that improved the quality of this study.

Author Contributions

Chengxi Huang : collection, analysis of the data, drafting the manuscript, finalising the writing.

Conflict of Interest

The author declares no conflict of interest in the present study, and no significant financial support for this work could have influenced its outcome.

Declaration of Artificial Intelligence (AI) Assistance

The author declares no generative AI and AI-assisted technologies in the process of writing the manuscript.

Ethics Approval

This work was not submitted to any other journal in any form, and the results of this study were not used in any animal experiments or human studies.

Funding

This research received no external funding.

References

1. National Library of Medicine. Traditional Chinese Medicine. In: National Library of Medicine. 2025. https://www.nlm.nih.gov/hmd/topics/chinese-traditional/index.html
2. Furhad S, Sina RE, Bokhari AA. Cupping Therapy. In: StatPearls. Treasure Island (FL): StatPearls Publishing. 2025. https://www.ncbi.nlm.nih.gov/books/NBK538253/
3. Chen X, Cui J, Li R, *et al*. Dao Yin (a.k.a. Qigong): Origin, development, potential mechanisms, and clinical applications. Evidence-Based Complementary and Alternative Medicine. 2019;2019(1): 3705120. https://onlinelibrary.wiley.com/doi/pdf/10.1155/2019/3705120
4. Wang Z, Wang D, Liu W, Wang Z. Traditional Chinese medicine diagnosis and treatment based on systematics. ILIVER. 2023;2:181-7.
5. Zhao C, Li GZ, Wang C, *et al*. Advances in patient classification for traditional Chinese medicine: A machine learning perspective. Evidence-Based Complementary and Alternative Medicine. 2015;2015(1):376716. https://onlinelibrary.wiley.com/doi/pdf/10.1155/2015/376716
6. Johns Hopkins Medicine. Chinese Medicine. In: Johns Hopkins Medicine. 2019. https://www.hopkinsmedicine.org/health/wellness-and-prevention/chinese-medicine
7. Gaur R. A brief history: Traditional Chinese medicinal system. Pharmacological Research - Modern Chinese Medicine. 2024;10: 100387. https://doi.org/10.1016/j.prmcm.2024.100387
8. World Health Organization. WHO international standard terminologies on traditional medicine in the Western Pacific region. World Health Organization. 2007;251 Regional Office for the Western Pacific. https://catalog.nlm.nih.gov/permalink/01NLM_INST/vdtut1/alma9913193053406676

9. World Health Organization. WHO traditional medicine strategy: 2014-2023. World Health Organization, 2023. https://www.who.int/publications/i/item/9789241506096
10. Wang J, Chen L, Wang Z, *et al*. Perceptions, attitudes, beliefs, and practices toward traditional Chinese medicine therapies for herpes zoster: A cross-sectional study of healthcare professionals. Risk Management and Healthcare Policy. 2024; 17:1239-51.
11. Huang Y, Zhou M, Deng Q, *et al*. Bibliometric analysis for the literature of traditional Chinese medicine in PubMed. Scientometrics. 2015; 105:557-66.
12. Consentino R, Santos MJ, Matos LC, *et al*. Bibliometric analysis of traditional Chinese medicine scientific production between 1982 and 2016 indexed in PubMed. Medicines. 2018;5(2):41. https://www.mdpi.com/2305-6320/5/2/41
13. Pan D, Guo Y, Fan Y, *et al*. Development and application of traditional Chinese medicine using AI machine learning and deep learning strategies. The American Journal of Chinese Medicine. 2024; 52:605-23.
14. Jongjiamdee K, Pornwonglert P, Na Bangchang N, *et al*. Artificial intelligence in traditional medicine: evidence, barriers, and a research roadmap for personalized care. Frontiers in Artificial Intelligence. 2025;8: 1659338. https://www.frontiersin.org/journals/artificial-intelligence/articles/10.3389/frai.2025.1659338/pdf
15. Aria M, Cuccurullo C. Bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics. 2017;11: 959-975.
16. Nees VE, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 2010;84: 523-538.
17. Jelodar H, Wang Y, Yuan C, *et al*. Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. Multimedia Tools and Applications. 2019;11: 15169-211.
18. OpenAI. ChatGPT. 2025. https://chatgpt.com
19. Bar-Ilan J. The h-index of h-index and of other informetric topics. Scientometrics. 2008; 75:591-605.
20. Sievert C, Shirley K. LDAvis: A method for visualizing and interpreting topics. In: Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces. 2014;63-70. https://aclanthology.org/W14-3110.pdf
21. Tian G, Zhao C, Zhang X, *et al*. Evidence-based traditional Chinese medicine research: Two decades of development, its impact, and breakthrough. Journal of Evidence-based Medicine. 2021; 14:65-74.
22. Zhang W, Ma L, Xie W, *et al*. Advances in the application of traditional Chinese medicine during the COVID-19 recovery period: A review. Medicine. 2024;103: e37683. https://journals.lww.com/md-journal/_layouts/15/oaks.journals/downloadpdf.aspx?an=00005792-202404050-00054
23. Lu DP, Lu GP. An historical review and perspective on the impact of acupuncture on U.S. medicine and society. Medical Acupuncture. 2013; 25:311-6.
24. Hao JJ, Mittelman M. Acupuncture: Past, present, and future. Global Advances in Health and Medicine. 2014; 3:6-8.
25. Fan X, Meng F, Wang D, *et al*. Perceptions of traditional Chinese medicine for chronic disease care and prevention: A cross-sectional study of Chinese hospital-based health care professionals. BMC Complementary and Alternative Medicine. 2018; 18:209. https://link.springer.com/content/pdf/10.1186/s12906-018-2273-y.pdf
26. Ji Z, Lee N, Frieske R, *et al*. Survey of hallucination in natural language generation. ACM Computing Surveys. 2023;12:1-38.
27. Nori H, King N, McKinney SM, *et al*. Capabilities of gpt-4 on medical challenge problems. ArXiv. 2023;13375. https://arxiv.org/pdf/2303.13375
28. Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. The New England journal of medicine. 2016 Sep 29;375(13):1216.
29. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Center Report. 2019;49:15-21. https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/hast.973?casa_token=CVJilEGHcqAAAAAA%3AiIlZiG8-4UIlk0B629WFmaQZQ2wuRRgSFGdwTG2LNlCbx4qax8qLv0BfiQANt0jn9awRAB-66dxdlx4v

How to Cite: Huang C. Too Early to Call? - A Critical Evaluation of AI's Nascent Role in Comprehension for Traditional Chinese Medicine. Int Res J Multidiscip Scope. 2026; 7(1): 1063-1075.
DOI: 10.47857/irjms.2026.v07i01.08145