Current Status Analysis of Outcome Measures in Randomized Controlled Trials of Traditional Chinese Medicine for Chronic Kidney Disease (Postprint)

Authors: Zhou Jiali, Ze Yuan, Yu Li, Du Yutong, Hao Na, Hao Na

Date: 2026-05-07T16:33:44+00:00

Abstract

Background: According to the Global Burden of Disease study, the prevalence of chronic kidney disease (CKD) is increasing annually. Traditional Chinese Medicine (TCM) has demonstrated unique advantages in delaying CKD progression and improving renal function; however, clinical research faces issues such as non-standardized outcome measures, which affect efficacy evaluation and evidence synthesis. Objective: To analyze the application status and current situation of outcome measures in randomized controlled trials (RCTs) of TCM for CKD, providing a reference for constructing a core outcome set (COS) for TCM treatment of CKD. Methods: A computer-based search was conducted in databases including CNKI, Wanfang Data, VIP, SinoMed, PubMed, Cochrane Library, and Web of Science for RCTs related to TCM treatment of CKD. The search period ranged from 2014-01-01 to 2024-10-31. The RoB 2.0 tool was used to assess the risk of bias, and statistical analysis was performed on the outcome measures. Results: A total of 298 RCTs were finally included, reporting 217 types of outcome measures with a total frequency of 2,539. The outcome measures were classified into 7 categories, ranked by frequency as follows: laboratory indicators (1,851 times, 72.90%), symptoms/signs (253 times, 9.96%), TCM syndromes (203 times, 8.00%), safety events (201 times, 7.92%), quality of life (22 times, 0.87%), long-term prognosis (9 times, 0.35%), and economic evaluation (1 time, 0.34%). Conclusion: The current status of outcome measures is characterized by a wide variety of indicators, non-standardized nomenclature, lack of distinction between primary and secondary outcomes, inconsistent efficacy evaluation criteria, significant variations in measurement time points, and neglect of quality of life and economic evaluation indicators. Due to the lack of unified standards for outcome measures, it is necessary to actively construct a core outcome set for TCM treatment of CKD to improve the quality of clin-

ical research design, enhance the reliability of clinical research evidence, and subsequently promote the high-quality development of TCM research.

Full Text

Preamble

Chinese General Practice

Abstract

In recent years, the rapid development of artificial intelligence and big data technologies has significantly influenced the field of medicine. As the foundation of the healthcare system, general practice faces unique challenges and opportunities in the era of intelligence. This article explores the integration of machine learning and deep learning within the framework of Chinese general practice, aiming to provide a comprehensive overview of current applications and future directions.

Introduction

General practice serves as the first point of contact for patients and plays a critical role in disease prevention, chronic disease management, and health promotion. However, the increasing complexity of clinical data and the shortage of primary care resources necessitate more efficient diagnostic and management tools. The emergence of machine learning offers a promising solution to these challenges by enabling the analysis of vast amounts of clinical information to support decision-making.

Applications of Machine Learning in Primary Care

The application of machine learning in general practice is multifaceted, ranging from early screening to personalized treatment plans. By utilizing algorithms such as random forests, support vector machines, and neural networks, researchers can identify high-risk populations for chronic conditions like hypertension and diabetes more accurately than traditional statistical methods.

1.1 Early Screening and Risk Prediction Early intervention is a cornerstone of general practice. Machine learning models can integrate electronic health records (EHRs), lifestyle factors, and genetic information to predict the likelihood of disease onset. For instance, predictive models for cardiovascular risk have shown improved sensitivity and specificity when incorporating nonlinear variables through deep learning techniques.

1.2 Chronic Disease Management Managing chronic diseases requires continuous monitoring and adjustment of treatment strategies. Intelligent systems can assist general practitioners (GPs) by analyzing patient-generated health

data from wearable devices. This allows for real-time feedback and alerts when physiological parameters, such as blood glucose or blood pressure, deviate from the target range.

[Figure 1: see original paper]

Challenges and Ethical Considerations

Despite the potential benefits, several hurdles remain in the widespread adoption of AI in Chinese general practice. Data privacy and security are paramount, as the training of robust models requires access to sensitive patient information. Furthermore, the "black box" nature of some deep learning models poses challenges for clinical interpretability, which is essential for building trust between doctors and patients.

The mathematical representation of model reliability can be expressed through the loss function $\mathcal{L}$, where we aim to minimize the discrepancy between predicted outcomes $\hat{y}$ and actual clinical observations y:

$$\mathcal{L} = \sum (y - \hat{y})^2$$

https://www.chinagp.net E-mail:zgqkyx@chinagp.net.cn

· Evidence-Based Medicine ·

Outcome Measures in Randomized Controlled Trials of Traditional Chinese Medicine for Chronic Kidney Disease Zhou Jiali[1,2], Yuan Ze[1,2], Li Yu[1,2], Du Yutong[1,2], Hao Na[1,2*]

Background

According to the Global Burden of Disease study, the prevalence of chronic kidney disease (CKD) is increasing annually. Traditional Chinese Medicine (TCM) has demonstrated significant clinical advantages in delaying the progression of CKD. However, the complex composition of TCM formulas and the multi-target nature of their mechanisms pose substantial challenges for pharmacological research.

In recent years, the integration of machine learning and deep learning technologies has provided new methodologies for exploring the scientific basis of TCM. By utilizing high-throughput data and computational modeling, researchers can more effectively identify active compounds, predict molecular targets, and elucidate the synergistic effects of herbal combinations in treating CKD. This interdisciplinary approach bridges the gap between traditional wisdom and modern evidence-based medicine, offering promising avenues for the development of novel therapeutic strategies.

Traditional Chinese Medicine (TCM) has demonstrated unique advantages in delaying the progression of Chronic Kidney Disease (CKD) and improving renal

function. However, clinical research in this field faces challenges such as the lack of standardized outcome measures, which hinders the objective evaluation of efficacy and the integration of high-quality evidence.

Objective: This study aims to analyze the current application and status of outcome measures in randomized controlled trials (RCTs) of TCM for CKD. The findings are intended to provide a reference for the development of a Core Outcome Set (COS) for TCM treatments in CKD.

We conducted a comprehensive computer-based search across several major academic databases, including the China National Knowledge Infrastructure (CNKI), Wanfang Data Knowledge Service Platform, VIP Database (CQVIP), China Biology Medicine (CBM), PubMed, and the Cochrane Library.

Methods

Data Sources and Search Strategy

A comprehensive search was conducted for randomized controlled trials (RCTs) concerning Traditional Chinese Medicine (TCM) for the treatment of Chronic Kidney Disease (CKD) across several major databases, including PubMed, Cochrane Library, and Web of Science. The search period spanned from January 1, 2014, to October 31, 2024.

Quality Assessment and Data Analysis

The Risk of Bias 2.0 (RoB 2.0) tool was employed to evaluate the methodological quality and risk of bias of the included studies. Statistical analysis was subsequently performed on the extracted data, focusing on primary and secondary outcome measures to synthesize the clinical evidence.

Results

Ultimately, 298 randomized controlled trials were included in the study, all of which reported outcome measures.

A total of 217 outcome measures were identified, with a cumulative frequency of 2,539 occurrences. These measures were categorized into seven distinct types. Ranked by frequency, they are: physical and chemical testing (1,851 occurrences, 72.90%), symptoms and signs (253 occurrences, 9.96%), Traditional Chinese Medicine (TCM) syndromes (203 occurrences, 8.00%), safety events (201 occurrences, 7.92%), quality of life (22 occurrences, 0.87%), long-term prognosis (9 occurrences, 0.35%), and economic evaluation (1 occurrence, 0.34%).

Conclusion

The current state of outcome measures is characterized by a vast diversity of types, a lack of standardized nomenclature, and an absence of clear prioritization

between primary and secondary endpoints.

To address issues such as the fragmentation of outcome measures, inconsistent efficacy evaluation standards, significant variations in measurement time points, and the neglect of quality-of-life and economic evaluation indicators, further action is required. The current lack of unified standards for outcome measures necessitates the active construction of a Core Outcome Set (COS) for Traditional Chinese Medicine (TCM) in the treatment of Chronic Kidney Disease (CKD). Establishing such a set will improve the quality of clinical research design and enhance the reliability of clinical evidence, thereby driving the high-quality development of TCM research.

Keywords: Chronic kidney disease; Traditional Chinese medicine; Randomized controlled trial; Outcome measures; Core outcome set

Abstract

Chronic Kidney Disease (CKD) is a major global public health challenge characterized by high prevalence, poor prognosis, and significant medical costs. Traditional Chinese Medicine (TCM) has demonstrated unique advantages and clinical efficacy in delaying the progression of CKD and improving patient quality of life. However, existing clinical research on TCM for CKD faces several critical issues, including high heterogeneity in outcome measures, a lack of patient-reported outcomes, and insufficient focus on long-term clinical endpoints. These deficiencies hinder the high-quality evaluation and international dissemination of TCM interventions.

The development of a Core Outcome Set (COS) provides a standardized solution to these challenges. By defining a minimum set of outcomes that should be measured and reported in all clinical trials for a specific health condition, a COS facilitates the comparison and synthesis of data across different studies. This paper reviews the current status of outcome measures in randomized controlled trials (RCTs) of TCM for CKD, analyzes existing problems, and discusses the necessity and methodological framework for establishing a TCM-specific COS for CKD. Such a framework should integrate the characteristics of TCM syndrome differentiation with international standards, ensuring that outcome selection is scientifically rigorous, clinically relevant, and representative of both clinician and patient perspectives.

Introduction

Chronic Kidney Disease (CKD) refers to chronic abnormalities in kidney structure or function lasting for more than three months. As a progressive disease, it often leads to end-stage renal disease (ESRD), requiring dialysis or transplantation. In recent years, the integration of Traditional Chinese Medicine (TCM) and Western medicine has become a cornerstone of CKD management in China. While numerous randomized controlled trials (RCTs) have been conducted to

validate the efficacy of TCM, the lack of standardized outcome measures remains a significant barrier to evidence-based practice.

Current Status of Outcome Measures in TCM RCTs for CKD

Current clinical trials of TCM for CKD typically rely on a variety of indicators, which can be broadly categorized into laboratory parameters, clinical endpoints, and TCM-specific symptoms.

Laboratory Parameters

Most studies focus on surrogate markers such as Serum Creatinine (SCr), Estimated Glomerular Filtration Rate (eGFR), and 24-hour urinary protein. While these are essential for monitoring disease progression, they do not always correlate directly with patient-centered outcomes or long-term survival.

Clinical Endpoints

Hard endpoints, such as the doubling of serum creatinine, progression to ESRD, or all-cause mortality, are the "gold standards" for evaluating treatment efficacy in CKD.

[Chinese Library Classification Number] R 692

Abstract

This paper presents a comprehensive study on the application of machine learning and deep learning techniques in modern scientific research. By integrating advanced computational models with traditional analytical frameworks, we demonstrate significant improvements in predictive accuracy and data processing efficiency. Our methodology emphasizes the importance of maintaining mathematical rigor while leveraging the scalability of neural networks. The results indicate that the proposed approach outperforms existing benchmarks across several key performance metrics, providing a robust foundation for future developments in the field.

Introduction

The rapid advancement of machine learning has fundamentally transformed the landscape of data-driven scientific inquiry. As datasets grow in complexity and volume, traditional statistical methods often struggle to capture high-dimensional non-linear relationships. Deep learning, a subset of machine learning characterized by multi-layered neural architectures, has emerged as a powerful tool for addressing these challenges. This study explores the synergy between theoretical modeling and empirical data analysis, aiming to bridge the gap between abstract mathematical formulations and practical implementation.

Methodology

3.1 Data Preprocessing and Feature Engineering

Before training the models, the raw data undergoes a rigorous preprocessing pipeline. This includes normalization, handling of missing values, and the extraction of relevant features. Let $\mathcal{X}$ represent the input space and $\mathcal{Y}$ the target space. The goal is to learn a mapping function $f : \mathcal{X} \to \mathcal{Y}$ that minimizes the empirical risk. We define the feature vector as $x_i \in \mathbb{R}^d$ and the corresponding label as y_i.

3.2 Model Architecture

The core of our approach involves a deep neural network architecture designed to handle sequential data. We utilize a combination of convolutional layers for spatial feature extraction and recurrent units for temporal dependencies. The loss function employed is the mean squared error (MSE), defined as:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^{N} \|y_i - \hat{y}_i\|^2$$

where $\hat{y}_i = f(x_i; \theta)$ is the predicted value and θ represents the model parameters. To prevent overfitting, we incorporate dropout layers and L_2 regularization.

[Figure 1: see original paper]

Results and Discussion

Our experimental results demonstrate that the proposed model achieves a high degree of precision. As shown in [Figure 2: see original paper], the convergence rate of the training process is significantly improved.

DOI: 10.12114/j.issn.1007-9572.2025.0257

Analysis of Current Status of Outcome Measures in Randomized Controlled Trials of Traditional Chinese Medicine for Chronic Kidney Disease ZHOU Jiali[1,2], YUAN Ze[1,2], LI Yu[1,2], DU Yutong[1,2], HAO Na[1,2*] 1. The First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin 300381, China 2. National Clinical Research Center for Chinese Medicine Acupuncture and Moxibustion, Tianjin 300381, China

Abstract Background: According to the Global Burden of Disease Study, the prevalence of chronic kidney disease (CKD) is increasing worldwide. Traditional Chinese Medicine (TCM) has demonstrated unique advantages in delaying CKD progression and improving renal function; however, clinical research in this field is hampered by non-standardized outcome measures, which impedes consistent efficacy evaluation and evidence synthesis. **Objective:** To analyze the application and current status of outcome measures in randomized controlled trials

(RCTs) of TCM for CKD, and to provide a reference for establishing a core outcome set (COS) for TCM interventions in CKD. **Methods:** RCTs on TCM for CKD published between January 1, 2014, and October 31, 2024, were retrieved from the following databases: CNKI, Wanfang Data, VIP, SinoMed, PubMed, Cochrane Library, and Web of Science. The risk of bias for the included studies was assessed using the RoB 2.0 tool.

Funding: (22KPHDRC00040)

Suggested Citation: ZHOU Jiali, YUAN Ze, LI Yu, et al. Analysis of current status of outcome measures in randomized controlled trials of traditional Chinese medicine for chronic kidney disease [J]. *Chinese General Practice*, 2025. DOI: 10.12114/j.issn.1007-9572.2025.0257. [Epub ahead of print] [www.chinagp.net]

https://www.chinagp.net E-mail:zgqkyx@chinagp.net.cn

Chinese General Practice

Results: A total of 298 RCTs were included, reporting 217 different outcome measures with a total frequency of 2,539. These outcomes were categorized into seven domains, in descending order of frequency: laboratory tests (1,851; 72.90%), symptoms/signs (253; 9.96%), TCM syndrome patterns (203; 8.00%), safety events (201; 7.92%), quality of life (22; 0.87%), long-term prognosis (9; 0.35%), and economic evaluation (1; 0.34%). **Conclusion:** Current outcome measures exhibit significant heterogeneity, non-standardized terminology, lack of distinction between primary and secondary outcomes, inconsistent efficacy criteria, variable assessment timepoints, and insufficient attention to quality of life and economic evaluations. There is an urgent need to develop a COS for TCM in CKD to improve the quality and reliability of clinical research, thereby promoting high-quality development of TCM evidence.

Key words: Chronic kidney disease; Traditional Chinese medicine; Randomized controlled trial; Outcome measures; Core outcome set

Chronic kidney disease (CKD) is a chronic condition characterized by structural damage and functional impairment of the kidneys (with a medical history ≥ 3 months) resulting from various factors [**?**]. According to the Global Burden of Disease (GBD) study, the global prevalence of CKD has reached 9.1%, while the prevalence in China is as high as 10.8%. In the same year, approximately 1.2 million deaths were attributed to CKD, and this figure is projected to rise to 2.2 million by 2040, making it the fifth leading cause of death worldwide [**?**]. Currently, the treatment of CKD primarily focuses on managing primary diseases and symptomatic care. However, pharmacological options for renal protection remain extremely limited, and most patients eventually progress to end-stage renal disease. In recent years, numerous randomized controlled trials (RCTs) have demonstrated that Traditional Chinese Medicine (TCM) offers unique advantages in treating CKD, particularly in reducing proteinuria, protecting renal function, delaying disease progression, mitigating adverse drug reactions, and

alleviating related complications [**?**]. Outcome measures are the most critical factors linking interventions to efficacy; appropriate outcome measures can enhance the consistency and feasibility of RCTs while reducing reporting bias [**?**]. At present, the outcome measures used in RCTs for TCM treatment of CKD lack uniformity, featuring a wide variety of types, non-standardized criteria, and significant discrepancies in selection, which hinders the objective evaluation of TCM efficacy. Therefore, this study reviews the application of outcome measures in TCM treatments for CKD from 2014 to 2024 to provide a reference for the construction of a core outcome set (COS).

Materials and Methods

1.1 Search Strategy

A comprehensive computer-based search was conducted across several databases, including China National Knowledge Infrastructure (CNKI), Wanfang Data Knowledge Service Platform, VIP Database (CQVIP), China Biology Medicine (CBM), PubMed, Cochrane Library, and Web of Science. The search period spanned from January 1, 2014, to October 31, 2024. Chinese search terms included "chronic kidney disease," "chronic renal disease," "chronic renal insufficiency," "chronic renal failure," "Traditional Chinese Medicine," "Chinese patent medicine," "Chinese herbal medicine," "integrated Chinese and Western medicine,""acupuncture and moxibustion,""acupoint application,""massage,""external treatment,""randomized controlled trial,"and "randomized."English search terms included "chronic kidney disease," "chronic renal insufficiency," "chronic renal failure," "renal interstitial fibrosis," "Chinese medicine," "TCM," "herbal medicine," "acupuncture," and "acupoint patching." The search strategy utilized a combination of subject headings (MeSH/Emtree) and free-text terms.

1.2 Inclusion Criteria

The inclusion criteria for the literature were as follows: the study design must be a randomized controlled trial (RCT); the language must be Chinese or English; and the publications must be from Chinese core journals or SCI-indexed journals. The study population had to consist of patients with a clear clinical diagnosis of CKD. While there were no restrictions on the conventional treatments used in the control group, the intervention in the experimental group had to involve TCM modalities, such as Chinese herbal medicine, Chinese patent medicine, TCM injections, acupuncture, moxibustion, acupoint application, or TCM enemas. Finally, the studies were required to report at least two outcome measures.

1.3 Exclusion Criteria

Literature with unclear diagnostic criteria for Chronic Kidney Disease (CKD); studies involving patients with major comorbidities such as coronary heart disease or malignant tumors; duplicate publications, single-author papers, and

studies with incomplete data or unavailable full text; non-Randomized Controlled Trials (RCTs), including reviews, theoretical discussions, animal experiments, systematic reviews, conference papers, dissertations, and scientific achievements; and studies where the research subjects are undergoing peritoneal dialysis or hemodialysis.

1.4 Literature Screening and Data Extraction

The literature retrieved from the databases was imported into NoteExpress software. After automatic deduplication, two researchers independently screened the literature according to the inclusion and exclusion criteria. An initial screening was conducted by reading the titles and abstracts, followed by a secondary screening through full-text review. Finally, the two researchers cross-checked the results to determine the final included studies; any disagreements were resolved by a third researcher. The extracted data were imported into Excel and included the following: basic information (article title, year, and authors); study subjects (sample size, CKD stage, diagnostic criteria, and TCM syndrome types); interventions; risk of bias assessment; and outcome measures.

1.5 Risk of Bias Assessment

The risk of bias for the included studies was assessed using the Cochrane Risk of Bias tool (RoB 2.0) [**?**]. The evaluation criteria encompassed several key domains: generation of the random sequence, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, completeness of outcome data, selective reporting bias, and other potential sources of bias. Each domain was categorized as having a "low risk of bias," "unclear risk of bias," or "high risk of bias."

1.6 Outcome Measures

In accordance with the "Technical Specifications for the Development of Core Outcome Sets for Clinical Trials of Traditional Chinese Medicine," the outcome measures for Traditional Chinese Medicine (TCM) treatments for Chronic Kidney Disease (CKD) were standardized and categorized into seven distinct domains. These categories include: TCM disease patterns and syndromes, symptoms and physical signs, laboratory and physical examinations, quality of life, long-term prognosis, economic evaluation, and safety events. Descriptive analysis was subsequently performed on the classification and frequency distribution of these outcome measures [**?**].

1.7 Statistical Analysis

Chinese General Practice

Abstract

In the context of the ongoing reform of the medical and health system, the development of general practice has become a cornerstone for achieving "Healthy China" goals. This paper examines the current status, challenges, and future directions of general practice in China. By analyzing the integration of primary care services with advanced medical technologies, we explore how machine learning and deep learning can enhance diagnostic accuracy and patient management in community health settings. Our findings suggest that a robust general practice system is essential for managing chronic diseases and optimizing resource allocation across the healthcare hierarchy.

Introduction

General practice, as the foundation of the primary healthcare system, plays a vital role in providing continuous, comprehensive, and coordinated care to individuals and families. In China, the transition from a hospital-centric model to a community-based primary care model has accelerated in recent years. This shift is driven by the increasing burden of chronic non-communicable diseases and an aging population, which require long-term management rather than episodic acute care.

The integration of modern technology into general practice is no longer optional but a necessity for improving service efficiency. As noted in recent literature [**?**], the application of artificial intelligence and big data analytics allows general practitioners (GPs) to make more informed clinical decisions. However, the implementation of these technologies must be balanced with the humanistic care that defines the core of general practice.

Current Development of General Practice in China

The development of general practice in China has undergone significant evolution. Currently, the "5+3" standardized residency training program serves as the primary pathway for cultivating qualified GPs. Despite these efforts, a shortage of high-quality talent remains a bottleneck. illustrates the growth of the GP workforce over the last decade, highlighting both the progress made and the remaining gap in meeting the target of five GPs per 10,000 residents.

Furthermore, the "gatekeeper" role of GPs is being strengthened through the family doctor contract service model. This model encourages residents to seek initial consultations at community health centers, thereby reducing the pressure on tertiary hospitals. However, the effectiveness of this system depends on the public' s trust in the clinical competence of primary care providers.

Technological Integration and Machine Learning

To address the challenges of diagnostic complexity in primary care, researchers are increasingly turning to machine learning. By utilizing large-scale electronic health records (EHRs), deep learning models can predict disease progression and identify high-risk patients.

https://www.chinagp.net E-mail:zgqkyx@chinagp.net.cn

To evaluate the impact of the risk of bias on outcome measures, the included studies were categorized into three groups based on their risk of bias assessment: (1) Low risk of bias group: all domains were rated as "low risk of bias" ; (2) High risk of bias group: at least one key domain was rated as "high risk of bias" ; (3) Unclear risk group: all domains were rated as "unclear risk of bias" or a combination of "low risk" and "unclear risk," with no domains rated as "high risk." Fisher' s exact test was employed to compare the distribution of outcome measure types across these different groups, with $P < 0.05$ considered statistically significant.

(3) Blinding of participants and personnel: Seven studies implemented double-blinding and were rated as low risk of bias; 291 studies did not provide sufficient details regarding blinding and were rated as unclear risk of bias. (4) Blinding of outcome assessment: Only one study implemented blinding for outcome assessors and was rated as low risk of bias; the remaining studies did not mention this and were rated as unclear risk of bias. (5) Incomplete outcome data: Three studies explicitly reported patient attrition but took no measures during statistical analysis, resulting in a rating of high risk of bias; the others were rated as unclear risk of bias. (6) Selective reporting and other biases: Other potential biases, such as funding support and conflicts of interest, were unknown and thus all were rated as unclear risk [**?**], as shown in .

Results

2.1 Literature Search Results

A total of 7,102 documents were retrieved through the initial search strategy. After applying the inclusion and exclusion criteria, 298 documents were ultimately selected for analysis. This final sample consists of 295 Chinese publications and 3 English publications. The detailed literature screening process is illustrated in Figure 1 [Figure 1: see original paper].

The initial database search yielded 7,102 documents, distributed as follows: China National Knowledge Infrastructure (CNKI, n=1,867), PubMed (n=102), VIP Database (n=967), Cochrane Library (n=314), Wanfang Data (n=1,485), Web of Science (n=48), and SinoMed (n=2,319). After removing 3,071 duplicate records, 4,031 documents remained. Following a preliminary screening of titles and abstracts, 3,446 documents were excluded for failing to meet the criteria, leaving 585 documents for full-text review. During the secondary screening

phase, an additional 287 documents were excluded for the following reasons: theoretical discussions (n=97), animal experiments (n=58), systematic reviews (n=89), ineligible study populations (n=19), ineligible interventions (n=19), retrospective studies (n=4), and clinical observations (n=1). Consequently, a total of 298 documents were included in the final analysis.

2.2 Risk of Bias Results

2.2.1 Risk of Bias Assessment The RoB 2.0 tool was employed for the assessment, and the results revealed several issues, including non-standard clinical trial designs and inadequate blinding procedures. The specific evaluation items are as follows:

(1) Random sequence generation: A total of 179 RCTs described specific methods for random allocation. Among these, 149 studies utilized the "random number table method" and were subsequently rated as having a low risk of bias. 107 papers mentioned the word "random" but did not describe the specific allocation method, and were rated as having an unclear risk of bias. Twelve studies used the order of clinical visits or the time of admission for grouping and were consequently rated as having a high risk of bias. (2) Allocation concealment: Two studies were rated as low risk of bias, as they utilized sealed envelopes for allocation concealment; 296 studies did not mention this process and were rated as having an unclear risk of bias. (3) Blinding of participants and personnel···

Risk of Bias Assessment of Included Studies [n(%)]

Item	Low risk of bias	Unclear risk of bias	High risk of bias
Generation of random sequences	179 (60.07)	107 (35.91)	12 (4.03)
Allocation concealment	2 (0.67)	296 (99.33)	0 (0.00)
Blinding of patients and investigators	7 (2.35)	291 (97.65)	0 (0.00)
Blinding of outcome assessors	1 (0.34)	297 (99.66)	0 (0.00)
Whether the outcome data is complete	295 (98.99)	0 (0.00)	3 (1.01)
Selective reporting bias	0 (0.00)	298 (100.00)	0 (0.00)
Other bias	0 (0.00)	298 (100.00)	0 (0.00)

2.2.2 Risk of Bias Analysis Based on the statistical results, the risk of bias across the included studies was categorized as follows: 0 studies were identified as having a low risk of bias; 12 studies (4.03%) were classified as having a high risk of bias; and the remaining 286 studies (95.97%) were categorized as having an unclear risk of bias. Due to the absence of a low-risk group, subsequent analysis compared the high-risk and unclear-risk groups.

We investigated whether there were intergroup differences in the distribution of outcome domains between the "high risk of bias" group and the "unclear risk of bias" group. Fisher' s exact test was employed to compare the distribution of outcome domains across these two groups. The results indicated that the difference in the distribution of outcome domains between the two groups was not statistically significant ($P = 0.739$), as shown in .

2.3 Basic Characteristics of Included Literature

Basic information statistics were conducted for the 298 included studies, covering sample size, Chronic Kidney Disease (CKD) staging, Western medicine diagnostic criteria, Traditional Chinese Medicine (TCM) diagnostic criteria, TCM syndrome types, control group treatments, and interventions.

2.3.1 Sample Size A total of 27,271 patients were included across the studies. Notably, only 9 studies (3.02%) reported performing sample size estimation, while the remaining studies did not provide any such estimation.

2.3.2 CKD Staging Specific CKD staging was reported in 273 studies. The most frequent stage was CKD stages 3–4 (69 studies, 25.27%), followed by stages 1–5 (47 studies, 17.22%), stages 3–5 (42 studies, 15.38%), and stage 3 (28 studies, 10.26%). Other reported stages included stages 1–2 (15 studies, 5.49%), stages 2–3 (15 studies, 5.49%), stage 5 (14 studies, 5.13%), and stages 2–4 (14 studies, 5.13%).

2.3.3 TCM Diagnostic Standards A total of 207 papers mentioned specific diagnostic standards. The most frequently used was the "Guiding Principles for Clinical Research of New Chinese Medicines (Trial)" (128 occurrences).

Comparison of Outcome Domain Distribution Between Different Risk Groups [n(%)]

Group	Laboratory Tests	Symptoms / Signs	TCM Syndrome	Safety Events	Quality of Life	Long-term Prognosis	Economic Evaluation
High Risk of Bias Group	10/12	12/12	12/12	1/12	0/12	0/12	0/12
Unclear risk of bias group	188 (65.73)	239 (83.57)	286 (100.00)	25 (8.74)	8 (2.80)	1 (0.35)	134 (46.85)

https://www.chinagp.net E-mail:zgqkyx@chinagp.net.cn

Among the diagnostic standards used, the *Guiding Principles for Clinical Research on New Chinese Medicines (Trial)* was the most frequently cited (128 times, 61.84%), followed by the *Criteria for Diagnosis, Syndrome Differentiation, and Efficacy Evaluation of Chronic Renal Failure (Trial Scheme)* (20 times, 9.66%). The top 10 standards are detailed in .

Top 10 TCM Diagnostic Standards

TCM Diagnostic Standards	Frequency (Count)	Percentage (%)
Guiding Principles for Clinical Research on New Chinese Medicines (Trial)	128	61.84
Criteria for Diagnosis, Syndrome Differentiation, and Efficacy Evaluation of Chronic Renal Failure	20	9.66
Guiding Principles for Clinical Research on New Chinese Medicines for Chronic Nephritis (Trial)	12	5.80
Guidelines for the Integrated Traditional Chinese and Western Medicine Diagnosis and Treatment of Chronic Renal Failure	11	5.31

TCM Diagnostic Standards	Frequency (Count)	Percentage (%)
Internal Medicine of Traditional Chinese Medicine	10	4.83
Clinic Terminology of Traditional Chinese Medical Diagnosis and Treatment—Diseases	8	3.86
TCM Diagnosis and Treatment Protocols for 95 Diseases in 22 Specialties	7	3.38
Guidelines for the Diagnosis and Treatment of Common Diseases in TCM Internal Medicine (2008)	5	2.42
Criteria for Diagnosis and Therapeutic Effect of TCM Diseases and Syndromes	4	1.93
TCM Clinical Pathways for 104 Diseases in 24 Specialties	2	0.97

2.3.4 Western Medicine Diagnostic Criteria In the Western medicine diagnostic criteria, 258 papers mentioned specific diagnostic criteria, primarily the KDOQI guidelines, *Clinical Practice Guidelines for Chronic Kidney Disease*, published by the National Kidney Foundation (cited 62 times, 24.03%), followed by the KDIGO guidelines (cited 42 times, 16.28%). The top 10 diagnostic criteria are detailed in .

Top 10 Western Medicine Diagnostic Criteria

Western Medicine Diagnostic Criteria	Frequency (n)	Percentage (%)
KDOQI: Clinical Practice Guidelines for Chronic Kidney Disease	62	24.03
KDIGO	42	16.28

Western Medicine Diagnostic Criteria	Frequency (n)	Percentage (%)
Clinical Practice Guidelines for Chronic Kidney Disease and Dialysis	35	13.57
Practical Internal Medicine	32	12.40
Summary of the Symposium on the Classification, Treatment, and Diagnostic Criteria of Primary Glomerular Diseases	26	10.08
Guidelines for the Screening, Diagnosis, Prevention, and Treatment of Chronic Kidney Disease	21	8.14
Clinical Practice Guidelines for the Evaluation and Management of Chronic Kidney Disease	18	6.98
Guidelines for the Diagnosis and Treatment of Chronic Renal Failure	12	4.65
Clinical Diagnosis and Treatment Guidelines: Nephrology Volume	10	3.88

2.3.5 TCM Syndrome Distribution A total of 218 papers mentioned TCM syndrome types, involving a total of 60 syndrome types. The top 10 most frequently occurring Traditional Chinese Medicine (TCM) syndromes are as follows: Kidney Deficiency and Blood Stasis syndrome (99 instances, 45.41%), Spleen-Kidney Qi Deficiency syndrome (27 instances, 12.39%), Spleen-Kidney Yang Deficiency syndrome (19 instances, 8.72%), Spleen-Kidney Qi Deficiency with Dampness-Turbidity and Blood Stasis syndrome (16 instances, 7.34%), Damp-Heat syndrome (15 instances, 6.88%), Blood Stasis syndrome (13 instances, 5.96%), Spleen-Kidney Deficiency with Turbid-Toxin and Stasis Obstruction syndrome (12 instances, 5.50%), Spleen-Kidney Deficiency with Dampness-Turbidity syndrome (11 instances, 5.05%), Dampness-Turbidity

syndrome (9 instances, 4.13%), and Spleen-Kidney Qi Deficiency with Dampness-Turbidity syndrome (9 instances, 4.13%).

2.3.6 Interventions Based on the statistics of the 298 included papers, the ranking of interventions is as follows: Traditional Chinese Medicine (TCM) decoctions (183 papers, 61.41%), proprietary Chinese medicines (39 papers, 13.09%), hospital-made preparations (11 papers, 3.69%), and TCM injections (8 papers, 2.68%). Regarding external therapies, the distribution includes TCM enemas (62 papers, 20.81%), acupoint application (4 papers, 1.34%), TCM iontophoresis in the renal area (3 papers, 1.01%), anal administration (2 papers, 0.67%), TCM medicated baths (2 papers, 0.67%), moxibustion (2 papers, 0.67%), acupuncture (1 paper, 0.34%), acupoint injection (1 paper, 0.34%), and hot herbal compress (1 paper, 0.34%). Conventional treatments primarily focused on lowering blood pressure, managing blood glucose levels, protecting renal function, and providing symptomatic relief.

2.4 Outcome Measures

2.4.1 Classification of Outcome Measures The 298 papers included 7 types of outcome measures. The study identified a total of 217 outcome measures, which were utilized across the included literature with a cumulative frequency of 2,539 instances. Serum creatinine was the most frequently reported outcome measure, appearing in 281 studies (94.30%). This was followed by blood urea nitrogen (261 instances, 87.58%), clinical efficacy (247 instances, 82.89%), glomerular filtration rate (129 instances, 43.29%), and 24-hour urinary protein quantification (109 instances, 36.58%).

Notably, among the 298 analyzed papers, only four explicitly distinguished between primary and secondary outcome measures. Furthermore, only one study incorporated economic evaluation indicators. The classification, specific nomenclature, and frequency of the outcome measures identified in the included literature are detailed in .

2.4.2 Course of Treatment and Measurement Time Points Among the included literature, 291 papers mentioned specific treatment durations. The most frequently utilized course was 12 weeks, appearing in 68 papers (23.37%). This was followed by 8-week courses in 61 papers (20.96%) and 4-week courses in 55 papers (18.90%).

In terms of the selection of measurement time points, the studies were primarily divided into two categories: those measuring before and after treatment, and those measuring before, during, and after treatment. Specifically, 260 papers (89.35%) conducted measurements before and after the intervention, while 31 papers (10.65%) measured outcome indicators before, during, and after the treatment period, as shown in .

Discussion

This study conducted a statistical analysis of outcome measures for the treatment of Chronic Kidney Disease (CKD) with Traditional Chinese Medicine (TCM). A total of 298 studies were ultimately included, reporting 217 distinct outcome measures with a cumulative frequency of 2,539 occurrences. The analysis revealed the following issues:

3.1 Non-standardized Trial Design and High Risk of Bias

Among the 298 included studies, 179 (60.07%) specified the method of random assignment, while 12 (4.03%) utilized methods such as order of visit or time of admission, leading to a high risk of bias. Regarding the implementation of blinding, only 7 studies (2.35%) reported blinding of both patients and trial personnel. This deficiency increases the risk of performance and detection bias; particularly when evaluating highly subjective indicators—such as Traditional Chinese Medicine (TCM) syndromes and symptom scores—the expectancy effects of researchers or patients can significantly influence the results. To standardize clinical trial design, non-random assignment must be strictly prohibited, and randomization should be conducted using methods such as random number tables. Furthermore, double-blinding should be the standard, and researchers should adhere to CONSORT reporting guidelines and pre-register their study protocols. Standardized randomized controlled trials are essential to ensure the scientific validity of research findings and to minimize the risk of bias [**?**].

3.2 Inconsistent TCM Syndrome Names and Standards

Among the 298 included studies, 218 (73.15%) reported on Traditional Chinese Medicine (TCM) syndromes.

Classification and Frequency of Outcome Measures in Included Literature

Domain	Outcome Measures	Frequency (n)	Percentage (%)
TCM Syndrome	TCM syndrome scores	158	53.02
	TCM symptom scores	21	7.05
	Therapeutic efficacy of TCM syndromes	15	5.03
	TCM syndrome ratings	5	1.68
	TCM symptom ratings	4	1.34
Symptoms/Signs	Total clinical effective rate	247	82.89

Domain	Outcome Measures	Frequency (n)	Percentage (%)
	Symptom Rating Scale	4	1.34
	Quantitative Rating Scale for TCM Symptoms	2	0.67
Laboratory Tests	Serum creatinine (Scr)	281	94.30
	Blood urea nitrogen (BUN)	261	87.58
	Estimated glomerular filtration rate (eGFR)	129	43.29
	Endogenous creatinine clearance rate (Ccr)	18	6.04
	β_2-microglobulin (β_2-MG)	15	5.03
	Serum retinol-binding protein (RBP)	12	4.03
	α_1-microglobulin (α_1-MG)	10	3.36
	24-hour urinary protein quantification	109	36.58
	Urinary albumin (ALB)	35	11.74
	Urinary β_2-microglobulin	28	9.40
	Urinary microal-bumin/creatinine ratio	25	8.39
	Renal function; Urinary total protein (TP); Urine; Urinary albumin excretion rate	22	7.38
	Urinary α_1-microglobulin	18	6.04

Domain	Outcome Measures	Frequency (n)	Percentage (%)
	Urine protein-to-creatinine ratio (UPCR)	15	5.03
	Red blood cell count per high-power field (RBC/HPF) in urine	12	4.03
	Waxy cast count per low-power field (LPF)	10	3.36
	Transforming growth factor-β_1 (TGF-β_1); Connective tissue growth factor (CTGF)	8	2.68
	Prothrombin time (PT)	15	5.03
	Thrombin time (TT)	12	4.03
	Activated partial thromboplastin time (APTT)	10	3.36
	D-dimer	8	2.68
	CD4+/CD8+	12	4.03
	Natural killer (NK) cells; Circulating immune complexes (CIC)	10	3.36
	Soluble intercellular adhesion molecule-1 (sICAM-1)	8	2.68
	Mononuclear cell Toll-like receptor 4 (TLR4)	5	1.68
	Urinary κ light chain	4	1.34
	Urinary λ light chain	4	1.34

Domain	Outcome Measures	Frequency (n)	Percentage (%)
	Leukotactin-1 (Lkn-1)	2	0.67
	Superoxide dismutase (SOD)	15	5.03
	Malondialdehyde (MDA)	12	4.03
	Glutathione peroxidase (GSHPx)	10	3.36
	Nuclear factor-κB p65 subunit (NF-κBp65)	8	2.68
	Total antioxidant capacity (TAC)	5	1.68
	Lipid peroxide (LPO); AGEs; HO-1	4	1.34
	Catalase (CAT)	2	0.67
	Nuclear factor erythroid 2-related factor 2 (Nrf2)	2	0.67
	p-IκBα	2	0.67
	Advanced oxidation protein products (AOPP)	2	0.67
	C-reactive protein (CRP)	35	11.74
	Type IV collagen (C-IV)	28	9.40
	Interleukin-6 (IL-6)	25	8.39
	Serum Klotho	22	7.38
	Tumor necrosis factor-α (TNF-α)	18	6.04
	Bone morphogenetic protein-7 (BMP-7)	15	5.03

Domain	Outcome Measures	Frequency (n)	Percentage (%)
	High-sensitivity C-reactive protein (hs-CRP)	12	4.03
	Hepatocyte growth factor (HGF)	10	3.36
	Interleukin-1 (IL-1)	8	2.68
	Serum hyaluronic acid (HA)	5	1.68
	Interleukin-2 (IL-2)	4	1.34
	Fibronectin (FN)	4	1.34
	Interleukin-10 (IL-10)	2	0.67
	Fibroblast growth factor 23 (FGF23)	2	0.67
	Interleukin-8 (IL-8)	2	0.67
	Laminin (LN)	2	0.67
	Interleukin-17 (IL-17)	2	0.67
	Procollagen type III (PC-III)	2	0.67
	Urinary interleukin-18 (IL-18)	2	0.67
	Stromal cell-derived factor-1 (SDF-1)	2	0.67
	Interleukin-1β (IL-1β)	2	0.67
	α-smooth muscle actin (α-SMA)	2	0.67
	Interferon-γ (IFN-γ)	2	0.67
	Plasma fibrinogen (FIB)	15	5.03
	White blood cell (WBC) count	12	4.03
	Thrombospondin-1 (TSP-1)	10	3.36

Domain	Outcome Measures	Frequency (n)	Percentage (%)
	Plasminogen activator inhibitor-1 (PAI-1)	8	2.68
	Serum endotoxin	5	1.68
	Uric acid (UA); Cystatin C (CysC)	109	36.58
	N-acetyl-β-D-glucosaminidase (NAG)	15	5.03
	Urinary Kidney Injury Molecule-1 (Kim-1) levels	12	4.03
	Neutrophil fraction	10	3.36
	Neutrophil-to-lymphocyte ratio (NLR)	8	2.68
	Interleukin-4 (IL-4)	5	1.68
	Interleukin-12p70 (IL-12P70)	4	1.34
	Tumor necrosis factor receptor 1 (TNFR1)	2	0.67
	p-NF-κBp65	2	0.67
	Serum amyloid A	2	0.67
	$ 1 −	2	0.67
	$acidglycoprotein\|2\|0.67\|\|\|Matrixmetalloproteinase-9(MMP-9)\|2\|0.67\|\|\|UrinaryvimentinmRNA;Urinarymicroalbumin\|2\|0.67\|\|\|Carbondioxidecombistimulatinghormone(TSH)\|5\|1.68\|\|\|Erythrocyteaggregation;Wholebloodviscosity\|15\|5.0densitylipoprotein(LDL)\|25\|8.39\|\|\|High-densitylipoprotein(HDL)\|22\|7.38\|\|\|ApolipoproteinA1(Apo-A1)\|18\|6.04\|\|\|ApolipoproteinB100(ApoB100)\|15\|5.03\|\|\|Lipoprotein(a)[Lp(a)]\|12\|4.03\|\|\|Serlactate\|2\|0.67\|\|\|L-FABP;CHI3L1\|2\|0.67\|\|\|Brainnatriureticpeptide(BNP)\|2\|0.67\|\|\|TIMP-1\|2\|0.67\|\|\|Intercellularadhesionmolecule-1(ICAM-1)\|2\|0.67\|\|\|Hematocrit\|2\|0.67\|\|\|FT3;FT4\|2\|0.67\|\|\|1,25(OH)2D3\|2\|0.67\|\|\|Bicarbonate(HC$		

Domain	Outcome Measures	Frequency (n)	Percentage (%)
	Fasting insulin (FINS)	2	0.67
	HOMA-IR	2	0.67
	VEGF	2	0.67
	ADMA	2	0.67
	Renal artery resistance index (RI)	15	5.03
	ASL imaging measurements	12	4.03
	Renal cortical blood flow velocity	10	3.36
	Peak systolic velocity (PSV)	8	2.68
	Renal artery internal diameter	5	1.68
	Serum Hcy, NO, and ET-1	4	1.34
	EDV and Vs (cm/s)	2	0.67
	Escherichia coli	15	5.03
	Streptococcus sanguinis	12	4.03
	Bifidobacterium longum	10	3.36
	Lactobacillus bulgaricus	8	2.68
	Fecal sIgA	5	1.68
	Vs of segmental renal artery (cm/s)	4	1.34
	Main renal artery RI	2	0.67
	Segmental renal artery RI	2	0.67
	Mean perfusion intensity	2	0.67
	Hemoglobin (Hb)	35	11.74
	CBC and RBC count	28	9.40
	Platelet count (PLT)	25	8.39

Domain	Outcome Measures	Frequency (n)	Percentage (%)
	Albumin (ALB)	22	7.38
	Alanine aminotransferase (ALT)	18	6.04
	Total serum protein	15	5.03
	Serum alkaline phosphatase (ALP)	12	4.03
	γ-glutamyl transpeptidase (γ-GT)	10	3.36
	PI3K	8	2.68
	Akt; PI3K-Akt	5	1.68
	BECN1	4	1.34
	LC3-II	2	0.67
	Triceps skinfold thickness (TSF)	2	0.67
	Arm circumference (AC)	2	0.67
	Arm muscle circumference (AMC)	2	0.67
	Aspartate aminotransferase (AST)	2	0.67
	Serum Leptin and Chemerin	2	0.67
Quality of Life	SF-36	15	5.03
	KDQOL	12	4.03
	SGA	10	3.36
	MIS	8	2.68
	MS Assessment	5	1.68
	WHOQOL-BREF Scale	4	1.34
	KDQOL-SF™	2	0.67
Economic Evaluation	Economic Evaluation	1	0.34

Measurement Time Points in Included Literature

Measurement Time Points	Frequency (n)	Percentage (%)
Weeks 0 and 12	68	23.37
Weeks 0 and 8	61	20.96
Weeks 0 and 4	55	18.90
Weeks 0 and 24	32	11.00
Weeks 0 and 2	18	6.19
Weeks 0 and 3	12	4.12
Weeks 0 and 16	10	3.44
0, 60 d	8	2.75
0, 48 weeks	6	2.06
0, 30 d	5	1.72
0, 45 d	4	1.37
0, 15 d	3	1.03
0, 9 weeks	2	0.69
0, 72 d	2	0.69
0, 6 weeks	2	0.69
0, 20 weeks	1	0.34
0, 20 d	1	0.34
0, 1 week	1	0.34
0, 10 d	1	0.34
Weeks 0, 4, 8, 12, 16, 20, and 24	8	2.75
Weeks 0, 4, 8, and 12	6	2.06
Weeks 0, 8, 16, and 24	5	1.72
Weeks 0, 4, and 8	4	1.37
Weeks 0, 12, and 24	3	1.03
Weeks 0, 24, and 48	2	0.69
Weeks 0, 6, and 12	1	0.34
Weeks 0, 2, 4, 12, and 24	1	0.34
Weeks 0, 4, 8, and 13	1	0.34
0, 15, 45 d	1	0.34
0, 4, 8, 12, 24, 36, and 48 weeks	1	0.34
0, 2, and 4 weeks	1	0.34
0, 30, 60 d	1	0.34
0, 2, 4, and 8 weeks	1	0.34

There are a total of 60 different Traditional Chinese Medicine (TCM) syndromes identified, reflecting a vast diversity and a lack of unified reference standards. Many of these syndromes share the same underlying meaning despite variations in terminology; for example, "Spleen-Kidney Deficiency" is expressed through various terms such as *pixu shenxu*, *pishen liangxu*, and *pishen buzu*. Furthermore, because the clinical course of Chronic Kidney Disease (CKD) is protracted and often involves multiple concurrent patterns—such as blood stasis, dampness-turbidity, turbid toxins, phlegm-turbidity, and the mutual entanglement of phlegm and stasis [**?**]—the descriptions of these syndromes become

increasingly complex. This complexity makes it difficult to classify hybrid patterns such as "Spleen-Kidney Qi Deficiency with Dampness-Turbidity and Blood Stasis," "Spleen-Kidney Deficiency with Turbid Toxin and Stasis Obstruction," or "Spleen-Kidney Deficiency with Dampness-Turbidity." Consequently, the author suggests that pathogenesis should serve as the core guiding principle to unify the nomenclature and standards for TCM syndromes in CKD. Simplifying syndrome classification in this manner would facilitate the development of outcome measure scales specific to TCM characteristics in CKD, thereby establishing a Chinese medicine-based characteristic efficacy evaluation system [**?**].

3.3 Lack of Distinction Between Primary and Secondary Outcome Measures

Among the 298 included studies, only 4 (1.3%) distinguished between primary and secondary outcome measures. This represents a significant deviation from Item 12a of the SPIRIT 2025 statement, which requires the pre-definition of primary outcomes. Failure to clearly identify a primary outcome results in a lack of justification for sample size calculations and increases the risk of Type I and Type II errors [**?**]. Among these four studies, two utilized 24-hour urinary protein as the primary outcome, while the other two used glomerular filtration rate (GFR) and serum creatinine (Scr), respectively.

International standards for clinical research emphasize that distinguishing between primary and secondary outcomes is essential for enhancing research transparency and the credibility of results [**?**]. These standards require that outcomes be objective and easily measurable. The primary indicator serves not only as the basis for sample size estimation but also represents the clinical event most directly relevant and concerning to patients, thereby best reflecting the efficacy of the trial.

https://www.chinagp.net E-mail:zgqkyx@chinagp.net.cn

The selection of appropriate primary endpoints is essential for achieving research objectives and ensuring accurate sample size estimation, as determined by the specific experimental goals and clinical significance of the study [**?**, **?**]. However, the level of attention paid to this issue within the field of Traditional Chinese Medicine (TCM) has lagged significantly behind other disciplines. It is therefore necessary to place greater emphasis on primary outcome measures in clinical controlled trials and to clearly distinguish between primary and secondary endpoints during the study design phase [**?**].

3.4 Inconsistent Clinical Efficacy Evaluation Standards

Among the 298 included studies, 247 (82.89%) reported clinical efficacy. The primary metric used was the "effective rate," which was largely evaluated based on changes in clinical symptoms, laboratory indicators, and Traditional Chinese Medicine (TCM) symptoms. Specifically, some studies defined "clinical cure" as the complete disappearance of TCM symptoms and signs, or a reduction in

syndrome scores of $\geq$ 90% (converting score changes into an efficacy index). Other studies defined "marked effect" as a reduction in TCM syndrome scores of $\geq$ 60% and a decrease in serum creatinine (Scr) of $\geq$ 20%. Due to the use of different reference standards for efficacy, discrepancies exist across studies regarding the classification of clinical cure, marked effect, and overall effectiveness [**?**]. This lack of uniformity in clinical efficacy standards leads to inaccuracies in efficacy data, which subsequently compromises the reliability of evaluation results. There is an urgent need to standardize clinical efficacy evaluation criteria within the construction of core outcome sets. Establishing core endpoints —such as renal replacement events, changes in estimated glomerular filtration rate (eGFR), and Scr levels—is essential to promote the development of TCM efficacy evaluation toward evidence-based and standardized practices.

3.5 High Variance in Outcome Measure Frequency

A total of 217 outcome measures were identified, with a cumulative frequency of 2,539 uses. Statistical analysis revealed significant disparities in the frequency of these indicators; for instance, "clinical efficacy"was utilized 247 times, whereas 98 specific outcome measures were used only once. This inconsistency makes it difficult to compare research results or conduct integrated meta-analyses, leading to a substantial waste of health resources [**?**]. Furthermore, there are significant discrepancies in treatment durations and a lack of frequent measurements during the intervention period. Specifically, 260 studies (89.35%) performed measurements only before and after treatment, failing to capture the progression trajectory of Chronic Kidney Disease (CKD) or the dynamic changes in therapeutic efficacy. The time span for indicator evaluation also varied widely, ranging from a minimum of 7 days to a maximum of 48 weeks. Consequently, the development of a Core Outcome Set (COS) should clearly distinguish between mandatory and optional indicators, select reliable outcome measures, establish dynamic monitoring standards, and standardize treatment durations and measurement time points to improve the rigor of clinical research design.

3.6 Limited Application of Quality of Life and Economic Indicators

CKD has a long clinical course that severely impacts patients' quality of life. Given that a decline in quality of life is itself a risk factor for the progression of kidney disease, greater emphasis should be placed on patient well-being by leveraging the advantages of Traditional Chinese Medicine (TCM) treatments. Among the included literature, only 26 studies (8.93%) conducted quality of life evaluations, utilizing a total of seven different assessment scales. Only three studies employed the Kidney Disease Quality of Life Short Form (KDQOL-SF™), a specialized instrument consisting of two parts: the first comprises the Short Form Health Survey (SF-36), while the second is a kidney disease-specific module designed to evaluate conditions unique to renal patients. Most of the literature failed to use standardized tools for quality of life assessment and did not undergo content validity verification. According to COSMIN standards, unver-

ified instruments may fail to accurately reflect the impact of TCM treatments on the quality of life of CKD patients [**?**]. Furthermore, of the 298 included papers, only one conducted an economic evaluation, indicating a profound lack of attention toward economic indicators.

Traditional Chinese Medicine (TCM) demonstrates significant efficacy in treating CKD by delaying the progression of renal dysfunction and improving clinical symptoms. Due to its low cost, TCM can significantly reduce the direct medical burden on patients by decreasing the reliance on Western medications, such as renin-angiotensin system (RAS) inhibitors, and lowering the incidence of related complications.

Chinese General Practice

Safety evaluation has been neglected in the current literature. Among the 298 included studies, only 138 (46.31%) reported safety events, primarily focusing on adverse reactions. Most of these studies lacked detailed descriptions, failing to mention specific clinical manifestations, frequency, or the timing of adverse reactions. Safety indicators are a vital component of clinical evaluation in Traditional Chinese Medicine (TCM) and are as important as efficacy evaluation [**?**]. The CONSORT 2010 statement stipulates that safety evaluation should be included as an outcome measure and requires detailed reporting of the timing, severity, frequency, and causality of adverse events relative to the intervention [**?**]. Although TCM treatments are widely used and generally associated with fewer adverse reactions, this does not imply absolute safety. Researchers should place greater emphasis on the occurrence of adverse events and fully consider safety evaluation indicators related to interventions.

3.7 Prospects and Limitations

Regarding the inclusion of studies, this research did not retrieve relevant RCT protocols from clinical trial registration platforms, nor did it conduct a comprehensive analysis of observational, retrospective, or systematic review studies. This study has certain limitations: it did not perform a horizontal comparison of outcome indicators, lacked a systematic statistical analysis of the measurement tools for each indicator, and failed to conduct a more in-depth exploratory analysis of the outcome measures.

Consequently, the aforementioned issues limit the scope of the results. Future research should involve a more comprehensive and in-depth analysis and summary to construct a Core Outcome Set (COS) for Chronic Kidney Disease (CKD) with TCM characteristics. This will facilitate the standardization of efficacy evaluation and improve the level of evidence for TCM treatments [**?**].

Conclusion

An analysis of the included studies revealed several issues: inconsistent CKD staging standards, non-standardized clinical research designs, and a generally

high risk of bias. Furthermore, there is a lack of unified standards for TCM syndrome types. Regarding outcome indicators, the distinction between primary and secondary indicators is unclear, efficacy evaluation standards are insufficient, and the selection of indicators and measurement time points varies significantly across studies. Simultaneously, there is a marked lack of attention to patient quality of life, economic indicators, and safety evaluations. A Core Outcome Set (COS) can improve the quality of clinical trials, making them more standardized and regulated, thereby reducing heterogeneity in outcome indicators and enhancing the reproducibility and comparability of research.

The development of a COS facilitates the generation of high-level evidence, such as systematic reviews and meta-analyses, reduces outcome reporting bias, and improves the quality and reliability of evidence [**?**]. Therefore, the authors propose the following recommendations for constructing a COS for TCM treatment of CKD: (1) Standardize clinical design by following the CONSORT statement, utilizing computer-generated randomization, sealed envelope allocation concealment, and double-blind designs to ensure data integrity and standardized pre-registration of studies. (2) Clearly label CKD stages with reference to the staging management principles of the KDIGO guidelines, and design stratified outcome indicators based on the characteristics of each stage. This will enhance the clinical applicability of evidence and lay the foundation for a COS with TCM characteristics. (3) Unify the nomenclature of TCM syndrome types according to the "Syndrome Naming Standards" in the appendix of the *Guiding Principles for Clinical Research of New Chinese Medicines*. Use the Delphi method to reach an *Expert Consensus on TCM Pattern Differentiation for CKD* to streamline syndrome types, improve the TCM efficacy evaluation system, and highlight the unique advantages of TCM. (4) Clearly define primary and secondary outcome indicators.

Chinese General Practice

Following the selection criteria for core indicators in the COMET handbook (primary indicators ≤ 2), primary indicators should be "hard endpoints," such as the rate of decline in eGFR. Furthermore, researchers should: (5) standardize clinical efficacy evaluation criteria and establish grading standards; (6) reasonably select outcome indicators and measurement time points; and (7) emphasize the assessment of quality of life, economic indicators, and safety events. This study summarizes and analyzes outcome indicators to provide a basis for the design of outcome measures in TCM research for CKD. Subsequent studies may refer to this research, as well as high-quality domestic and international guidelines and systematic reviews, to commit to the construction of a COS for TCM treatment of CKD, thereby enhancing the quality and reference value of clinical research.

Author Contributions: Zhou Jiali proposed the main research objectives and was responsible for the conceptualization, design, implementation, and drafting of the paper. Yuan Ze, Du Yutong, and Li Yu were responsible for data collection, organization, and the creation and presentation of figures and tables. Hao

Na was responsible for quality control and revision, overall accountability for the article, and supervision and management.

The authors declare no conflicts of interest. Zhou Jiali https://orcid.org/0009-0004-6500-726X

References

https://www.chinagp.net E-mail:zgqkyx@chinagp.net.cn

[1] KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease [J]. Kidney Int, 2024, 105 (4S): S117-S314. DOI: 10.1016/j.kint.2023.10.018. [2] CHRONIC KIDNEY DISEASE COLLABORATION G B D. Global, regional, and national burden of chronic kidney disease, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017 [J]. Lancet, 2020, 395 (10225): 709-733. DOI: 10.1016/S0140-6736(20)30045-3. [3] JAGER K J, KOVESDY C, LANGHAM R, et al. A single number for advocacy and communication—worldwide more than 850 million individuals have kidney diseases [J]. Kidney Int, 2019, 96 (5): 1048-1050. DOI: 10.1016/j.kint.2019.07.012. [4] WANG G R, AN X Y, SONG L Q. Current status of network pharmacology in the treatment of chronic kidney disease with traditional Chinese medicine [J]. Tianjin Journal of Traditional Chinese Medicine, 2020, 37 (8): 955-960. [5] QIU R J, SUN Y, HU J Y, et al. Methods for selecting outcome measurement instruments in clinical research [J]. Chinese Journal of Evidence-Based Medicine, 2018, 18 (2): 238-243. [6] GU H Q, WANG Y, LI W. Application of the Cochrane risk of bias tool in Meta-analysis of randomized controlled trials [J]. Chinese Circulation Journal, 2014, 29 (2). [7] STERNE J A C, SAVOVIĆ J, PAGE M J, et al. RoB 2: a revised tool for

Note: Figure translations are in progress. See original paper for figures.