Original Article

Using PRECIS-2 in Chinese herbal medicine randomized controlled trials for irritable bowel syndrome: A methodological exploration based on literature

Minjing Luo [a,b], Yingqiao Wang [b], Jinghan Huang [b], Yilin Li [c], Wenjie Li [b], He Li [b], Zhihan Liu [a,b], Meijun Liu [a,b], Yunci Tao [a,b], Jianping Liu [a,b], Yutong Fei [a,b,*]

[a] *Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China*
[b] *School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China*
[c] *School of Qi-Huang Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China*

ARTICLE INFO

ABSTRACT

Background: The pragmatism levels of randomized controlled trials (RCTs) mean how similar the interventions delivered in the trial setting match those in the setting where the results will be applied. We aimed to investigate the association between the consistency of pragmatism among the characteristics of RCT design and its effect size of results in Chinese herbal medicine (CHM) for irritable bowel syndrome (IBS).

Methods: Eight English and Chinese language databases were searched for RCTs on CHM for IBS. Six reviewers independently assessed the pragmatism of trials using the pragmatic-explanatory continuum indicator summary 2 (PRECIS-2) tool. The consistency of pragmatism levels among the characteristics of RCT design was calculated using the coefficient of variation. Linear regression models were adopted to explore influence factors of the pragmatism of RCTs.

Results: 78 RCTs were included. The level of consistency in the pragmatism for RCT's design was significantly correlated with the effect size of the results (binary outcome, r = -0.413; $P = 0.005$; continuous outcome, r = -0.779, $P < 0.001$). PRECIS-2 score was higher in trials with individualized interventions than fixed interventions (3.29 [0.32] vs 2.90 [0.32]; Cohen's d relative effect size, 0.52; P < 0.001) and in standard or usual-treatment-controlled trials than placebo-controlled (3.05 [0.37] vs 2.83 [0.28]; Cohen's d relative effect size, 0.32; P = 0.048).

Conclusion: The consistency of pragmatism level across the 9 domains of the PRECIS-2 tool in CHM IBS RCTs was positively correlated with the effect size of the results.

1. Introduction

Randomized controlled trials (RCTs) are considered the gold standard for evaluating intervention effects due to their rigorous designs, which balance known and unknown confounding factors.[1] Since Schwartz and Lellouch first elaborated the concept of explanatory and pragmatic design in RCTs in 1967,[2,3] there described two purposes for RCTs. A pragmatic RCT is conducted in real-world settings and involves usual care, aiming to inform decisions about whether to implement an intervention. An explanatory RCT is carried out in a controlled, idealized environment to maximize the chances of demonstrating the intervention's beneficial effects. The pragmatism level of an RCT means how similar the intervention delivered in the setting in which the trial was conducted and the intervention delivered in the setting in which its results are applied, which plays a crucial role in clinical decision-making.[4]

To help trialists comprehend the level of pragmatism of RCTs and make design decisions that serve the intended purpose of the trial in an even better fashion, the Pragmatic Explanatory Continuum Index Summary (PRECIS) tool was developed in 2008[5] and formed the updated PRECIS-2 version in 2015.[6]

Several studies[7-9] have used the PRECIS-2 tool to assess the pragmatism level of trials, illustrating how current trials can help researchers gain confidence in applying the studied interventions in the real-world. In RCTs that evaluated the effects of Chinese herbal medicine (CHM) interventions, unique challenges arise when assessing pragmatism. These trials often use subjective traditional Chinese medicine (TCM) syndrome

---

diagnostic criteria as part of their eligibility criteria[10] and employ complex interventions that combine standardized principal CHM formulations with personalized modifications based on symptoms within the framework of TCM theory.[11-13] The unique characteristics of CHM may complicate the design decisions of such RCTs and make significant heterogeneity in the pragmatism levels across the different domains of the PRECIS-2 tool. However, there is a paucity of data on how the consistency of pragmatism levels across the 9 domains of the PRECIS-2 tool has influenced the effect size of results in RCTs.

Irritable bowel syndrome (IBS) is one of the most common chronic digestive disorders, which is characterized by abdominal pain and discomfort, defecation as well as change in stool consistency and frequency.[14] TCM has been used to treat symptoms associated with IBS for thousands of years.[15] Plenty of previous RCTs, including explanatory and pragmatic RCTs, have evaluated the effect and safety of CHM formulae in the treatment of IBS.[11,12,16,17] This research included RCTs comparing CHM treatments for IBS as an example, aiming to investigate the relevance of consistency among RCT design characteristics to the effect size of the results. It also examined potential associations between specific study characteristics and the level of pragmatism of the RCTs to demonstrate the importance of consistent design decisions for a more accurate and scientific evaluation of CHM trial outcomes.

2. Methods

2.1. Literature search

To identify potentially eligible studies, we conducted a comprehensive search across Medline (Ovid), Embase, Cochrane Library, Web of Science, China National Knowledge Infrastructure (CNKI), SinoMed, Chinese Scientific Journal Database, and Wan-Fang databases from their inception until January 31, 2023. Our search strategy utilized a combination of keywords related to "Chinese medicine", "irritable bowel syndrome", "randomization", and "clinical trials". We imposed no language restrictions and also examined the reference lists of relevant studies for additional reports. The full search strategy can be found in Supplement 1.

2.2. Study selection

2.2.1. Types of trials
We included controlled trials that explicitly utilized random allocation to treatment.

2.2.2. Types of participants
Trials were conducted among adult patients with IBS encompassing various subtypes such as diarrhea-predominant, constipation-predominant, mixed, or other types.

2.2.3. Types of interventions
The treatment included single herbs, Chinese proprietary herbal medicines, and Chinese herbal formulas. There were no restrictions on the formulation of herbal compounds or the incorporation of integrative medicine. Studies involving non-oral administration modes for herbal medicines were excluded.

2.2.4. Comparison group
The control group included placebo, treatment as usual, no treatment, or positive interventions.

2.2.5. Outcomes
The primary outcomes included the effective rate and response rate of the IBS Symptom Severity Scale (IBS-SSS),[18] the total score on the IBS-SSS scale,[18] the adequate relief (IBS-AR),[19] the response rate of abdominal pain measured on the visual analogue scale (VAS scale), the abdominal pain score on the VAS scale,[20] the response rate on the Bristol Stool Scale,[21] and the score on the IBS Quality of Life questionnaire.[19]

2.3. Data collection process

Four researchers (Li YL, Wang YQ, Huang JH, and Liu ZH) used the Excel software to extract data from the final included RCTs in a standardized format. The extracted data included the study title, publication year, sample size, funding status, pathological type of IBS, inclusion and exclusion criteria, recruitment details, trial settings (single center or multiple centers), organizational details, intervention flexibility level (whether the interventions were tailored to individual patient conditions), types of interventions compared (placebo, standard treatment, treatment as usual, or no treatment), adherence enhancement strategies, follow-up duration, primary outcome, and whether an intention-to-treat analysis was conducted (yes, or no). The Cochrane Collaboration's risk of bias tool 2.0 (RoB 2.0) was employed to assess the risk of bias for each included RCT.[22]

2.4. PRECIS-2 tool

The level of pragmatism for each included trial was assessed using the PRECIS-2 tool, which included nine domains of trial design (eligibility criteria, recruitment, setting, organization, the flexibility of intervention delivery, the flexibility of adherence to the intervention, follow-up, primary outcome, and primary analysis),[23] with a scale of 1–5 to each domain (1 = maximal explanatory, 3 = equally pragmatic and explanatory, 5 = maximal pragmatic).[6] The assessment was first conducted independently by six raters with diverse backgrounds in TCM and clinical epidemiology, and the results were transmitted from each rater to Luo MJ, who then performed the statistical analysis. A consensus meeting was held among all raters to discuss discrepancies until an agreement was achieved. An RCT-specific summary PRECIS-2 score was calculated by averaging the scores over the nine domains, referred to as the mean PRECIS-2 score. The coefficient of variation (CV),[24] a statistical metric used to gauge the relative dispersion of data points around the mean in a data series, was used to evaluate the consistency of the pragmatism levels across the scoring results of the nine domains of the PRECIS-2 tool. A lower CV value indicates a higher consistency level of pragmatism among the characteristics of the RCT. Intraclass correlations were employed to evaluate the level of agreement among raters (inter-rater reliability) both before and after the consensus conference.[25]

2.5. Statistical analysis

Categorical variables were summarized using frequency and percentages, while continuous variables were summarized using mean and standard deviation (SD). The relative risk (RR) was calculated to quantify the effect size for binary outcomes, whereas Cohen's d was calculated for continuous outcomes.[26] Interpretation of changes or differences of continuous outcomes between groups was categorized as small, medium, or large based on Cohen's d values of 0.2 to 0.49, 0.5 to 0.79, and 0.8 or more, respectively. The changes or differences of binary outcomes between groups were interpreted as small, medium, or large based on RR values of 1.0 to 1.4, 1.5 to 2.9, and 3.0 to 9.9, respectively.[26] The Cohen's d was also used to quantify the mean difference between RCTs with different characteristics in relation to the level of pragmatism. Linear regression models were conducted to assess the associations between various trial characteristics and the level of pragmatism. Pearson correlation tested the trend between the RR or Cohen's d values and the CV for included trials. Statistical significance was considered at a two-sided $P \leq 0.05$. Data analysis was conducted using STATA version 15.0 software. A sensitivity analysis that explored the association between study characteristics and level of pragmatism was conducted based on a mean score of pragmatism level calculated by excluding domains that lacked of information (recruitment, organization, and the flexibility of adherence to the intervention).
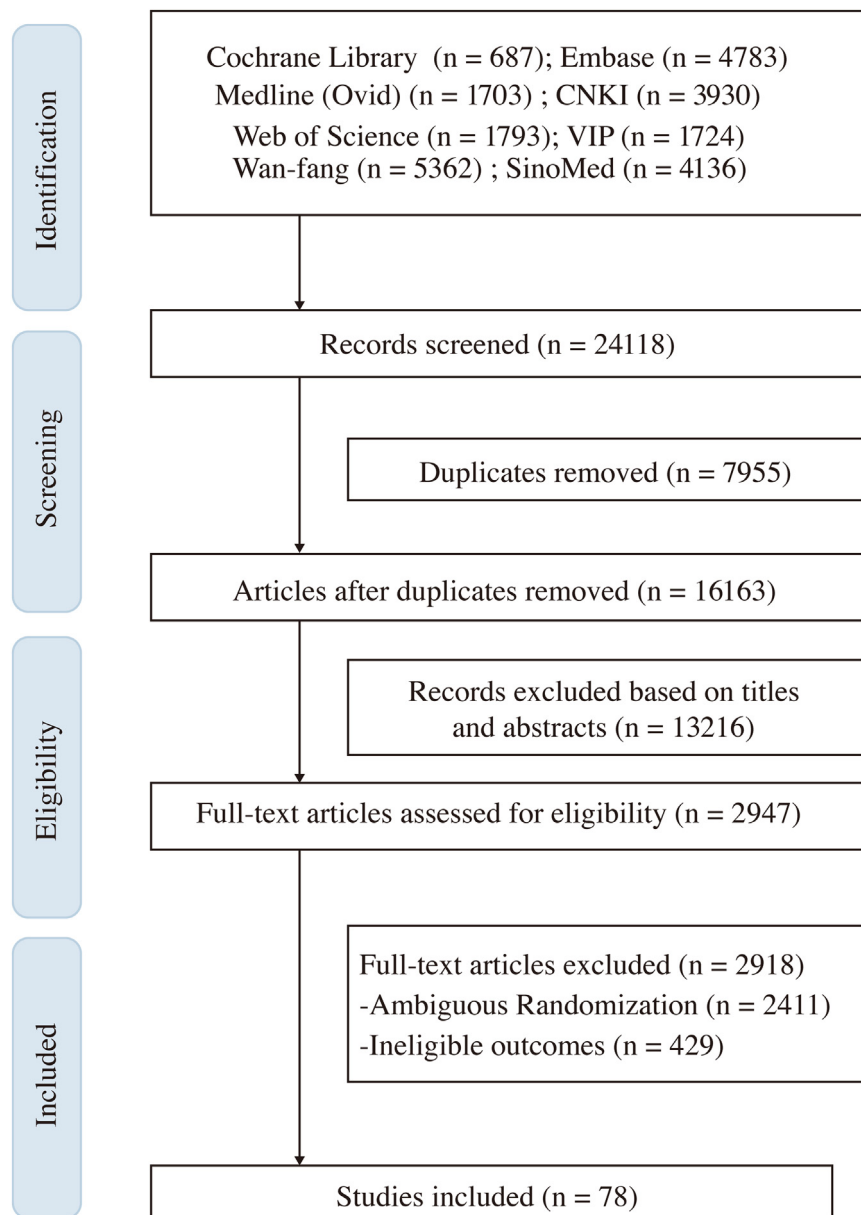
Fig. 1. Flow diagram of the literature screening process.

3. Results

Out of a total of 24,118 citations initially identified as potentially relevant, 78 RCTs were deemed eligible for inclusion. The selection process for all articles is presented in Fig. 1.

3.1. Characteristics of included trials

None of the included trials were explicitly self-identified by authors as explanatory or pragmatic RCTs. Characteristics of the 78 selected RCTs are summarized in Table 1. The majority of the RCTs were conducted in a single center (63, 80.77%), with 65 (83.33%) conducted compared to non-placebo-controlled. Pathological types were predominantly diarrhea in 64 RCTs (82.05%), followed by constipation only in 6 (7.69%) and mixed in 8 (10.26%) RCTs. Concerning the implementation of TCM syndrome differentiation, over three-quarters (59, 75.64%) of the eligible trials included patients' TCM syndromes as criteria for inclusion. However, less than one-third of the trials (22, 28.21%) were conducted with interventions that flexibly tailored to the patient's condition. The primary outcome was classified as a binary variable in 45

(57.69%) studies and as a continuous variable in 33 (42.31%) studies. Overall, 65 (83.33%) trials reported positive results for their primary outcome, while 13 (16.67%) were identified as negative. Based on the quality assessment of included RCTs, 14 (17.95%) trials were at low risk of bias, 14 (17.95%) were at high risk, and the remaining 50 (64.10%) had some concerns regarding the risk of bias.

3.2. PRECIS-2 scores

The mean (SD) PRECIS-2 score among the 78 RCTs was 3.11 (0.36), means the overall pragmatism level of included trials was neutral between full explanatory and full pragmatic. The setting and primary outcome domains received the lowest and highest PRECIS-2 ratings, respectively. Missing information in publications was frequently cited as the main reason for the discrepancy in initial scoring among raters (intra-class correlation coefficient, ICC = 0.569).[27] Improved inter-rater reliability was observed after the consensus discussion (ICC = 0.901). Although full agreement among raters was not achieved, the maximum difference was 1 point.

Table 1
Study Characteristics and Level of Pragmatism.

Factors	No. (%)	Score, Mean (SD)	Effect Size	P-value
Overall	78 (100)	3.11 (0.36)	NA	NA
Year of publication				
~2010	8 (10.26)	2.91 (0.46)	[Reference]	0.587
2011-2015	12 (15.38)	3.16 (0.44)	0.26	
2016-2020	35 (44.87)	2.93 (0.32)	0.03	
2021-2023	23 (29.49)	3.09 (0.32)	0.22	
Pathological type				
IBS-M	8 (10.26)	2.89 (0.36)	[Reference]	0.249
IBS-D	64 (82.05)	3.02 (0.37)	0.18	
IBS-C	6 (7.69)	3.12 (0.35)	0.31	
TCM differentiation				
No	19 (24.36)	2.99 (0.37)	[Reference]	0.795
Yes	59 (75.64)	3.01 (0.35)	0.03	
Treatment flexibility				
Fixed	56 (71.79)	2.90 (0.32)	[Reference]	**< 0.001**
Flexible	22 (28.21)	3.29 (0.32)	0.52	
Placebo-controlled				
No	65 (83.33)	3.05 (0.37)	[Reference]	**0.048**
Yes	13 (0.17)	2.83 (0.28)	0.32	
Primary outcome type				
Continuous	33 (42.31)	2.99 (0.27)	[Reference]	0.726
Binary	45 (57.69)	3.02 (0.42)	0.04	
Primary outcome				
Negative	13 (0.17)	2.91 (0.32)	[Reference]	0.751
Positive	65 (83.33)	3.03 (0.37)	0.17	
No. of sites				
Single	63 (80.77)	3.06 (0.41)	[Reference]	0.058
Multicenter	15 (19.23)	3.26 (0.31)	0.27	
Risk of bias				
High	14 (17.95)	3.03 (0.34)	[Reference]	0.751
Some concerns	50 (64.10)	2.99 (0.38)	0.06	
Low	14 (17.95)	3.04 (0.32)	0.02	

3.3. Associations between study characteristics and the level of pragmatism

Table 1 shows the associations between study characteristics and the pragmatism levels of included trials. PRECIS-2 mean scores were higher in trials with interventions modified according to individualized symptoms than those with fixed interventions (mean [SD], 3.29 [0.32] vs 2.90 [0.32]; Cohen's d relative effect size, 0.52; $P < 0.001$) and in standard controlled or treatment-as-usual controlled trials than placebo-controlled (3.05 [0.37] vs 2.83 [0.28]; Cohen's d relative effect size, 0.32; $P = 0.048$). There was no significant association between the level of pragmatism and year of publication, pathological type (mixed, diarrhea, or constipation), inclusion criteria involving TCM syndrome differentiation for patients, type of primary outcome variable, or the positivity of primary outcome results. Although trials involving multiple sites had numerically higher PRECIS-2 mean scores than those conducted at a single site, the difference was not significant (3.26 [0.31] vs 3.06 [0.41]; $P = 0.058$). A similar trend was found in the sensitivity analysis (Supplement 2).

3.4. Relevance of the consistency in the pragmatism for RCT's design to its results

The consistency of pragmatism level across the nine domains of the PRECIS-2 tool in CHM IBS RCTs was significantly correlated with the Cohen's d for the continuous outcome (r = -0.779, $P < 0.001$) and weakly correlated with the RR value for the binary outcome (r = -0.413; $P = 0.005$, Fig. 2).

4. Discussion

4.1. Summary of findings

This study revealed a tendency towards larger effect sizes as the increased consistency of the pragmatism levels across the nine domains of

the PRECIS-2 tool, no matter whether the outcome was a binary or continuous variable (binary, r = -0.413; $P = 0.005$; continuous, r = -0.779, $P < 0.001$). Trials with individualized intervention or placebo-control were more significantly pragmatic than those with fixed intervention or non-placebo control.

4.2. Strengths and limitations

To the best of our knowledge, this study is the first to investigate the correlation between the consistency of pragmatism levels across the nine domains of the PRECIS-2 tool in each trial and the effect size of the results. Additionally, we conducted an extensive literature review. We explored the association between the level of pragmatism and various characteristics of trials, such as the type of intervention comparisons, the type of the primary outcome variable, the pathological type of IBS, and the assessment of the risk of bias. We also considered the unique features of RCTs in TCM, including the use of TCM syndrome differentiation as inclusion criteria for patients and the implementation of individualized treatments based on patients' symptoms. As a limitation, we did not examine the consistency between the trial's intended purpose and design decisions and how this consistency relates to the effect size of the results since none of the included trials were explicitly labeled as explanatory or pragmatic RCTs, and we could not categorize them into two extremes as there is no universally recognized critical trial feature for the classification. Furthermore, the lack of information on several domains of the PRECIS-2 tool in publications may introduce bias to the scoring process.

4.3. Comparison with previous studies

Recently, researchers have been increasingly using the PRECIS-2 tool to assess the pragmatism level of RCTs based on publications.[28-31] In line with the viewpoint of Merrick and his colleagues,[32] our research did not reveal any correlation between the settings and the level of pragmatism. We cannot simply assume that an RCT conducted at a single
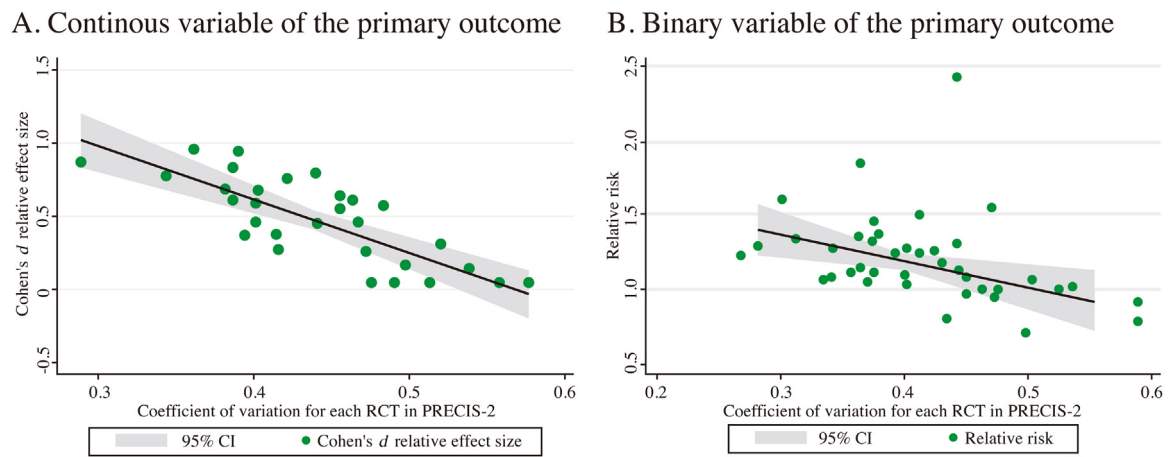
Fig. 2. Correlation trajectory between the coefficient of variation among the rating results on the 9 domains of PRECIS-2 tool in each trial and the effect size of each trial's result.

center is purely explanatory. A single-center RCT can offer valuable decision support for decision-makers in that specific center and similar ones.[33] Additionally, we propose another possible interpretation. Due to the widespread acceptance of TCM theory in China, there are not only Western hospitals but also specialized Chinese medicine hospitals and comprehensive hospitals that combine Chinese and Western medicine. Therefore, when assessing the "setting" domain, we considered not only the number of settings but also the suitability of Chinese medicine intervention procedures in Western or comprehensive hospitals.

Based on the recommendations of the PRECIS-2 team,[34] it becomes evident that certain crucial aspects of a trial, such as placebo control, may have minimal influence on the overall score. Devos et al. conducted a study involving 333 trials and found that the blinding of interventions did not result in any significant difference in the pragmatism levels of RCTs.[35] Similarly, another assessment conducted by Rafael revealed a consistent trend, suggesting that the PRECIS-2 tool may not effectively distinguish placebo-controlled trials from others.[36] However, a previous study[37] focused on cardiovascular clinical trials demonstrated that trials that controlled with device, behavioral interventions, or health system interventions had higher PRECIS-2 scores compared to trials using a placebo control (3.36 [0.70] vs 3.11 [0.66]; $P < 0.001$), which aligns with our findings (PRECIS-2 mean score, 3.05 [0.37] vs 2.83 [0.28]; $P = 0.048$). The unique characteristics of interventions, such as surgery, behavioral intervention, or TCM decoction, pose challenges in placebo intervention implementation.[38] Further research is needed to ascertain whether there is a correlation between the choice of comparator and the PRECIS-2 score.

4.4. Implication of future research

PRECIS-2 is not intended to design an explanatory or pragmatic trial.[39] Instead, it serves as a consensus process to assist research teams in carefully considering and assessing the level of pragmatism of the trial design.[6] The significant correlation between the consistency of the pragmatism levels across the nine domains in the PRECIs-2 tool and the effect size of the results indicates that when designing a research protocol, it is crucial to not only focus on the alignment between the purpose and design and the pragmatism level of each domain but also consider the consistency of the level of pragmatism among all nine domains of design characteristics. This process may facilitate achieving more positive results while ensuring the implementation of a trial design that aligns with the researcher's intended purposes.

Assessing the criteria of the nine domains was often challenging during the scoring process. Lipman and his colleagues pointed out that researchers' clinical expertise may impact scoring in each domain of

PRECIS-2.[23] Besides, unique features might exist for scoring each domain on the PRECIS-2 tool in the background of CHM treatment. Witt CM observed that much of the heterogeneity among the researchers was due to challenges in operationalizing the criteria, especially when the trial did not report any information on some domains.[23] Lu et al. found that researchers had different understandings and judgments about the usual care situation. Raters were uncertain about assessing the similarity of pragmatism levels between interventions delivered in the trial and those in "real world" treatment environments.[40] Moreover, scoring was uncertain regarding the lack of information in published articles, such as the recruitment process, the organization or training information for study participants, and the flexibility of adherence to the intervention. Furthermore, the sensitivity analysis excluding three domains lacking information showed lager PRECIS-2 mean score differences between trials with individualized interventions and those with fixed interventions, as well as trials conducted in multicenter and those in a single center.

In line with those challenges observed, we held a consensus meeting to discuss each discrepancy thoroughly and ultimately reached an agreement. Given the dynamic nature of individualized diagnosis and treatment in TCM,[41] many trials do not solely rely on a fixed intervention scheme for treatment. Instead, they may employ a fixed main prescription with modifications based on patient symptoms or entirely tailor the treatment based on patient symptoms. We may assume that both approaches are pragmatic, but their level of pragmatism varies.

In conclusion, the consistency of pragmatism level across the 9 domains of the PRECIS-2 tool in CHM IBS RCTs was positively correlated with the effect size of the results.

CRediT authorship contribution statement

Minjing Luo: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft. **Yingqiao Wang:** Methodology, Data curation. **Jinghan Huang:** Methodology, Data curation, Writing – review & editing. **Wenjie Li:** Data curation. **He Li:** Data curation. **Zhihan Liu:** Formal analysis, Data curation, Writing – original draft. **Meijun Liu:** Formal analysis, Data curation. **Yunci Tao:** Formal analysis, Data curation. **Jianping Liu:** Conceptualization. **Yutong Fei:** Conceptualization, Writing – review & editing.

Declaration of competing interest

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

Ethical statement

No ethical approval was required as this study did not involve human participants or laboratory animals.

Data availability

All data in our current study came from public databases, all results were presented in this article.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.imr.2024.101053.

References

1. Audrey S. Qualitative research in evidence-based medicine: improving decision-making and participation in randomized controlled trials of cancer treatments. *Palliat Med.* 2011;25(8):758–765.
2. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis.* 1967;20(8):637–648.
3. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Clin Epidemiol.* 2009;62(5):499–505.
4. Usman MS, Van Spall HGC, Greene SJ, et al. The need for increased pragmatism in cardiovascular clinical trials. *Nat Rev Cardiol.* 2022;19(11):737–750.
5. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol.* 2009;62(5):464–475.
6. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ.* 2015;350:h2147.
7. Bratton DJ, Nunn AJ. Alternative approaches to tuberculosis treatment evaluation: the role of pragmatic trials. *Int J Tuberc Lung Dis.* 2011;15(4):440–446.
8. Sharma T, Qamar I, Zwarenstein M. How pragmatic are randomized trials of remdesivir and favipiravir for in-hospital treatment of COVID-19: a descriptive methodological review of trial design using the PRECIS-2 framework. *J Clin Epidemiol.* 2022;152:193–200.
9. Witt CM, Manheimer E, Hammerschlag R, et al. How well do randomized trials inform decision making: systematic review using comparative effectiveness research measures on acupuncture for back pain. *PLoS One.* 2012;7(2):e32399.
10. Song J, Yang L, Su S, et al. The Diagnosis Performance of the TCM Syndromes of Irritable Bowel Syndrome by Gastroenterologists Based on Modified Simple Criteria Compared to TCM Practitioners: A Prospective. *Multicenter Preliminary Study. Evid Based Complement Alternat Med.* 2020;2020:9507674.
11. Shih YS, Tsai CH, Li TC, et al. The effect of Xiang-Sha-Liu-Jun-Zi tang (XSLJZT) on irritable bowel syndrome: A randomized, double-blind, placebo-controlled trial. *J Ethnopharmacol.* 2019;238:111889.
12. Tang XD, Lu B, Li ZH, et al. Therapeutic Effect of Chang'an I Recipe (I) on Irritable Bowel Syndrome with Diarrhea: A Multicenter Randomized Double-Blind Placebo–Controlled Clinical Trial. *Chin J Integr Med.* 2018;24(9):645–652.
13. Bensoussan A, Talley NJ, Hing M, Menzies R, Guo A, Ngu M. Treatment of irritable bowel syndrome with Chinese herbal medicine: a randomized controlled trial. *JAMA.* 1998;280(18):1585–1589.
14. Choung RS, Locke 3rd GR. Epidemiology of IBS. *Gastroenterol Clin North Am.* 2011;40(1):1–10.
15. Xiao HT, Zhong L, Tsang SW, Lin ZS, Bian ZX. Traditional Chinese medicine formulas for irritable bowel syndrome: from ancient wisdoms to scientific understandings. *Am J Chin Med.* 2015;43(1):1–23.
16. Li CY, Ain Mohd Tahir N, Li SC. A systematic review of integrated traditional Chinese and Western medicine for managing irritable bowel syndrome. *Am J Chin Med.* 2015;43(3):385–406.
17. Li DY, Dai YK, Zhang YZ, et al. Systematic review and meta-analysis of traditional Chinese medicine in the treatment of constipation-predominant irritable bowel syndrome. *PLoS One.* 2017;12(12):e0189491.
18. Francis CY, Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. *Aliment Pharmacol Ther.* 1997;11(2):395–402.
19. Bian LQ, Lu F, Li ZH, et al. [Analysis of Response of IBS-SSS, AR, and IBS-QOL in IBS Clinical Effect Evaluation]. *Zhongguo Zhong Xi Yi Jie He Za Zhi.* 2016;36(10):1191–1196.
20. Portincasa P, Bonfrate L, Scribano ML, et al. Curcumin and Fennel Essential Oil Improve Symptoms and Quality of Life in Patients with Irritable Bowel Syndrome. *J Gastrointestin Liver Dis.* 2016;25(2):151–157.
21. Drossman DA, Camilleri M, Mayer EA, Whitehead WE. AGA technical review on irritable bowel syndrome. *Gastroenterology.* 2002;123(6):2108–2131.
22. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 2019;366:l4898.
23. Lipman PD, Loudon K, Dluzak L, Moloney R, Messner D, Stoney CM. Framing the conversation: use of PRECIS-2 ratings to advance understanding of pragmatic trial design domains. *Trials.* 2017;18(1):532.
24. Pélabon C, Hilde CH, Einum S, Gamelon M. On the use of the coefficient of variation to quantify and compare trait variation. *Evol Lett.* 2020;4(3):180–188.
25. Kovačić J, Varnai VM. Intraclass correlation coefficient for grouped data. *Epidemiology.* 2014;25(5):769–770.
26. Cohen J. *Statistical Power Analysis for Behavioral Sciences.* Hillside, NJ: Lawrence Erlbaum Associates; 1988.
27. Dal-Ré R. Articles provided insufficient information to conduct an appropriate retrospective assessment of the pragmatic/explanatory features of medicine trials with the PRECIS-2 tool. *Eur J Clin Pharmacol.* 2020;76(8):1093–1102.
28. Johnson KE, Neta G, Dember LM, et al. Use of PRECIS ratings in the National Institutes of Health (NIH) Health Care Systems Research Collaboratory. *Trials.* 2016;17:32.
29. Glasgow RE, Gaglio B, Bennett G, et al. Applying the PRECIS criteria to describe three effectiveness trials of weight loss in obese patients with comorbid conditions. *Health Serv Res.* 2012;47(3 Pt 1):1051–1067.
30. Luoma KA, Leavitt IM, Marrs JC, et al. How can clinical practices pragmatically increase physical activity for patients with type 2 diabetes? A systematic review. *Transl Behav Med.* 2017;7(4):751–772.
31. Koppenaal T, Linmans J, Knottnerus JA, Spigt M. Pragmatic vs. explanatory: an adaptation of the PRECIS tool helps to judge the applicability of systematic reviews for daily practice. *J Clin Epidemiol.* 2011;64(10):1095–1101.
32. Zwarenstein M, Treweek S, Loudon K. PRECIS-2 helps researchers design more applicable RCTs while CONSORT Extension for Pragmatic Trials helps knowledge users decide whether to apply them. *J Clin Epidemiol.* 2017;84:27–29.
33. Zwarenstein M, Thorpe K, Treweek S, Loudon K. PRECIS-2 for retrospective assessment of RCTs in systematic reviews. *J Clin Epidemiol.* 2020;126:202–206.
34. Forbes G, Loudon K, Treweek S, Taylor SJC, Eldridge S. Understanding the applicability of results from primary care trials: lessons learned from applying PRECIS-2. *J Clin Epidemiol.* 2017;90:119–126.
35. Devos F, Foissac F, Bouazza N, Ancel PY, Tréluyer JM, Chappuy H. Study characteristics impacted the pragmatism of randomized controlled trial published in nursing: a meta-epidemiological study. *J Clin Epidemiol.* 2019;116:18–25.
36. Dal-Ré R. The PRECIS-2 tool seems not to be useful to discriminate the degree of pragmatism of medicine masked trials from that of open-label trials. *Eur J Clin Pharmacol.* 2021;77(4):539–546.
37. Sepehrvand N, Alemayehu W, Das D, et al. Trends in the Explanatory or Pragmatic Nature of Cardiovascular Clinical Trials Over 2 Decades. *JAMA Cardiol.* 2019;4(11):1122–1128.
38. Pardo-Cabello AJ, Manzano-Gamero V, Puche-Cañas E. Placebo: a brief updated review. *Naunyn Schmiedebergs Arch Pharmacol.* 2022;395(11):1343–1356.
39. Loudon K, Zwarenstein M, Sullivan FM, et al. The PRECIS-2 tool has good interrater reliability and modest discriminant validity. *J Clin Epidemiol.* 2017;88:113–121.
40. Lu L, Zhou L, Dong J, Xiang Y, Wen Z. The application of PRECIS-2 ratings in randomized controlled trials of Chinese herbal medicine. *Oncotarget.* 2017;8(63):107002–107010.
41. Liu S, Zhu JJ, Li JC. The interpretation of human body in traditional Chinese medicine and its influence on the characteristics of TCM theory. *Anat Rec (Hoboken).* 2021;304(11):2559–2565.