Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 12 July 2026
Revised 02 Aug 2026
Accepted 15 Aug 2026

Natural Resources for Human Health 2026, Vol. 6 (18s): 670–689

KEYWORDS:
Premature ovarian insufficiency; Oral Chinese herbal medicine; Traditional Chinese medicine; Hormone replacement therapy; Systematic review; Meta-analysis

Efficacy and Safety of Oral Traditional Herbal Medicine in the Treatment of Premature Ovarian Insufficiency: A Systematic Review and Meta-Analysis

Vu Thi Thuy Linh[1], Vu Ha My[2], Le Minh Tien[3], Nguyen Duc Manh[4]

[1] Dai Nam University, Hanoi, Vietnam
[2] VNU University of Medicine and Pharmacy, Hanoi, Vietnam
[3] Hanoi Medical University Hospital, Hanoi, Vietnam
[4] Nghe An Traditional Medicine Hospital, Nghe An, Vietnam

ABSTRACT: Premature ovarian insufficiency (POI) is a condition characterized by impaired ovarian function before the age of 40 years. Oral traditional herbal medicine is used either alone or in combination with hormone replacement therapy (HRT); however, its efficacy and safety remain unclear. This study evaluated traditional herbal medicine versus HRT, and traditional herbal medicine plus HRT versus HRT. Five databases were searched up to 27 July 2026. Randomized controlled trials (RCTs) were analyzed using a random-effects model and stratified by treatment duration of ≤3 months and >3 months. Risk of bias and certainty of evidence were assessed using the Cochrane Risk of Bias 2 (RoB 2) tool and the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach, respectively. A total of 28 RCTs were included. Traditional herbal medicine alone yielded inconsistent results, whereas traditional herbal medicine combined with HRT showed more pronounced improvements in several biological markers. In the ≤3-month subgroup, the combination of traditional herbal medicine and HRT significantly reduced follicle-stimulating hormone (FSH) levels (SMD = −1.61; 95% CI: −1.96 to −1.26) and increased estradiol (E2) levels (SMD = 1.07; 95% CI: 0.71–1.43), anti-Müllerian hormone (AMH) levels (SMD = 1.35; 95% CI: 0.59–2.11), and antral follicle count (AFC) (MD = 1.17; 95% CI: 0.78–1.56). The clinical effective rate was also higher with traditional herbal medicine plus HRT. No clear differences in adverse events were observed between treatment groups. However, many analyses showed substantial heterogeneity, and the certainty of the evidence was predominantly low or very low. Traditional herbal medicine, particularly when combined with HRT, may improve certain biological markers in patients with POI; nevertheless, further high-quality RCTs with long-term reproductive outcomes are warranted.

1. INTRODUCTION

Premature ovarian insufficiency, also known as premature ovarian failure, is a clinical syndrome characterized by impaired ovarian function in women before the age of 40 years(1). The condition typically presents with menstrual disturbances accompanied by elevated gonadotropin levels and reduced estrogen levels. A meta-analysis published in 2023 reported that the global prevalence of POI among women is as high as 3.5%(2). Notably, the prevalence of POI was significantly higher in developing countries than in developed countries (5.3% vs. 3.1%, respectively), with the highest prevalence reported in North America, reaching 11.3% (2). In recent years, the prevalence POI among women younger than 35 years has reached as high as

4%, with an increasing trend over time and a growing proportion of cases occurring at younger ages (3). POI not only has a profound impact on reproductive health but also adversely affects women's physical and psychological well-being. Affected women may experience amenorrhea, infertility, osteoporosis, cardiovascular disease, and other complications, thereby significantly impairing their quality of life.

Current treatment of POI mainly includes hormone therapy, assisted reproductive technologies, and lifestyle modifications. (4). These approaches may improve ovarian function to some extent; however, their effectiveness in restoring reproductive function remains limited (5). In addition, serious adverse events may occur during treatment, including an increased risk of endometrial cancer, breast cancer, and stroke …(1). Therefore, there is an urgent need to identify safe and effective adjunctive treatments with fewer adverse effects.

According to traditional Chinese medicine (TCM), premature ovarian insufficiency does not have an exact corresponding disease entity. Based on its clinical manifestations, however, it may be classified under TCM syndromes such as amenorrhea, blood dryness, scanty menstruation, and infertility(6). The pathogenesis is primarily attributed to kidney deficiency and premature depletion of Tiangui. Prolonged stagnation of liver qi may further impair the heart and spleen, thereby reducing the generation and transformation of qi and blood and ultimately resulting in amenorrhea. The main treatment principles include nourishing kidney essence, soothing the liver and relieving qi stagnation, and tonifying the heart and spleen. Traditional Chinese herbal medicine is prescribed according to different therapeutic principles, primarily targeting the underlying pathophysiological patterns of kidney deficiency, liver qi stagnation, and concurrent heart-spleen deficiency. Numerous clinical studies have evaluated the effects of traditional Chinese herbal medicine on menstrual recovery, hormone levels, and ovarian function in patients with premature ovarian insufficiency. Although several systematic reviews and meta-analyses have assessed traditional Chinese herbal medicine for the treatment of POI, the scope of previous studies remains fragmented. Wu et al. (2026)(7) focused on oral traditional Chinese herbal medicine alone compared with HRT and did not evaluate the adjunctive role of traditional Chinese herbal medicine when combined with HRT. Regarding combination therapy, Zhong et al. (2022)(8) onducted a Bayesian network meta-analysis of proprietary Chinese medicines combined with HRT; however, the analysis was limited to 12 proprietary Chinese medicine preparations, and the literature search was conducted only up to July 2022. More recently, Lin et al. (2025)(9) evaluated Kuntai capsules combined with HRT; however, the review focused on a single proprietary Chinese medicine preparation, and the literature search included studies published only up to October 2020. Therefore, previous reviews have not comprehensively evaluated the two commonly used strategies—oral traditional Chinese herbal medicine alone and traditional Chinese herbal medicine combined with HRT—within the same analytical framework, nor have they covered the full range of oral herbal formulations used in clinical practice. The present study was conducted to separately synthesize evidence comparing TCM with HRT and TCM plus HRT with HRT, while incorporating more recent RCTs. Analyses were further stratified by treatment duration (≤3 months and >3 months) to assess the consistency of treatment effects over time. In addition to endocrine and ovarian reserve markers, the study also synthesized clinical effectiveness and adverse events and assessed the certainty of evidence using the GRADE approach, thereby providing a more comprehensive evaluation of the efficacy and safety of these two traditional Chinese herbal medicine strategies for POI.

2. METHODS

2.1. Protocol Registration and Reporting Guidelines

This systematic review and meta-analysis was registered in the International Prospective Register of Systematic Reviews (PROSPERO; registration number: CRD420261452890). The study was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines.(10). The study protocol and related registration information were publicly available in the PROSPERO database

The research question was formulated according to the PICOS framework. The study population comprised women younger than 40 years who were diagnosed with premature ovarian insufficiency (POI/POF). The intervention consisted of oral traditional Chinese medicine, administered either alone or in combination with HRT, whereas the control group received HRT alone. Eligible studies were required to report both FSH and E2 levels and at least one of the following ovarian reserve markers: AMH or AFC. Secondary outcomes included LH, overall clinical effectiveness, and adverse events. Eligible studies included randomized controlled trials (RCTs) and controlled clinical trials reported as using random allocation.

2.2. Data Sources and Search Strategy

We systematically searched PubMed, the Cochrane Library, China National Knowledge Infrastructure (CNKI), Wanfang Data, and the VIP Database from database inception to 27 July 2026. The final database search was completed on July 27, 2026, three days earlier than the planned end date specified in the PROSPERO protocol (July 30, 2026).

The search strategy combined controlled vocabulary terms, including Medical Subject Headings (MeSH) where applicable, with free-text terms related to premature ovarian insufficiency and traditional Chinese herbal medicine, using the Boolean operators AND and OR. The search strategy was adapted to the syntax and search fields of each database. Only studies published in English or Chinese were considered. The complete search strategies for all databases are provided in Appendix 1. Records retrieved from the databases were pooled and duplicates were removed before screening. The main search terms included: *"Premature Ovarian Insufficiency", "Premature Ovarian Failure", "Primary Ovarian Insufficiency", "Chinese Herbal Medicine", "Chinese herbal formula", "herb*", "decoction**". The search syntax was adapted to each database. For several Chinese-language databases, the corresponding Chinese terms were used, including: "早发性卵巢功能不全", "卵巢早衰", "原发性卵巢功能不全", "中药", "中草药", "中药复方", "中药方剂", "中药汤剂", "中药治疗".

2.3. Inclusion

Studies were included if they were randomized controlled clinical trials involving patients with premature ovarian insufficiency. POI was diagnosed according to the 2016 ESHRE criteria. For studies published before 2016, a diagnosis of POI/POF based on recognized professional criteria available at the time of the study was accepted, provided that the study population had ovarian dysfunction occurring before 40 years of age. In all cases, the diagnostic criteria had to be clearly described or referenced to a verifiable professional standard or guideline. The intervention group received oral traditional Chinese medicine, either alone or in combination with HRT, whereas the control group received HRT alone. No restrictions were imposed on herbal formula, composition, or dosage form, provided that the treatment was administered orally and that the ingredients, dosage, and administration were clearly described. Eligible studies were required to assess both outcome domains defined as central to this review: ovarian endocrine function and ovarian reserve. Accordingly, studies had to report FSH and E2 and at least one ovarian reserve marker, either AMH or AFC. This criterion was selected to ensure that the included studies provided evidence on both hormonal changes and ovarian reserve, in line with the primary objective of the review. In the registered PROSPERO protocol, FSH, E2, and AMH were prespecified as the core outcomes. During the review process, AFC was accepted as an alternative marker of ovarian reserve when AMH was not reported. This modification was documented as a deviation from the registered protocol.

2.4. Exclusion

Studies were excluded if they included additional treatment modalities such as acupuncture, moxibustion, enemas, injections, topical herbal applications, or other non-oral therapies. In vitro studies, animal studies, case reports, review articles, clinical experience reports, studies for which the full text was unavailable, studies with insufficient data for analysis, and studies that did not report both FSH and E2 together with at least one of the two ovarian reserve markers, AMH or AFC, were also excluded. Only full-text articles published in peer-reviewed journals were considered. When a thesis had subsequently been published as a journal article, only the corresponding article was included. If multiple reports used the same or overlapping study populations, only the most complete and relevant report was retained to avoid duplication of data.

2.5. Study Selection and Data Extraction

After duplicate records were removed, two reviewers (Vu Ha My and Le Minh Tien) independently screened the titles and abstracts according to the predefined inclusion and exclusion criteria. Potentially eligible studies were then assessed by full-text review. Any disagreements during the study selection process were resolved through discussion; if consensus could not be reached, a third reviewer (Vu Thi Thuy Linh) was consulted to make the final decision. When multiple reports included the same or overlapping study populations, the reports were cross-checked, and only the most complete or most appropriate report for each outcome was used to avoid duplication of data.

Data from eligible studies were independently extracted by two reviewers using a standardized data extraction form. The extracted information included author, year of publication, sample size, participant characteristics, age, disease duration,

diagnostic criteria, characteristics of the TCM intervention and control groups, herbal formula composition and dosage, HRT regimen, treatment duration, assessment time points, FSH, E2, AMH, AFC, LH, clinical effectiveness, and adverse events. For continuous outcomes, the number of participants, post-treatment mean values, and standard deviations were extracted. For dichotomous outcomes, the number of events and the total number of participants in each group were recorded. After extraction, the data were cross-checked against the original reports and between the two reviewers. Any discrepancies were resolved through discussion or consultation with a third reviewer.

When data were reported as medians and interquartile ranges, the studies were excluded from the primary analysis and were included only in sensitivity analyses after conversion to means and standard deviations, where appropriate. Data that were not reported or could not be reliably determined were recorded as not reported (NR) and were not imputed. Owing to limitations in contact information and the likelihood of obtaining responses, study authors were not contacted to obtain missing data; unreported data were recorded as NR

2.6. Outcome Assessment

The primary outcomes included FSH and E2, which were used to assess ovarian endocrine function, and AMH and AFC, which were used to reflect ovarian reserve. Secondary outcomes included overall clinical effectiveness. As no universally accepted criteria are currently available for assessing clinical effectiveness, efficacy measures based on recognized criteria used in clinical practice were considered acceptable. The total effective rate was calculated as follows: Total effective rate = cure rate + markedly effective rate + effective rate + improvement rate

2.7. Risk of Bias

To assess the risk of bias in the included studies, we used the Cochrane RoB 2 tool (11). The five RoB 2 domains assessed were: 1- bias arising from the randomization process; 2- bias due to deviations from intended interventions; 3- bias due to missing outcome data; 4- bias in measurement of the outcome; and 5- bias in selection of the reported result. Each domain, as well as the overall risk of bias, was classified as "low risk," "some concerns," or "high risk" according to the RoB 2 algorithm. Two reviewers independently assessed the risk of bias and compared their judgments. Any disagreements were resolved through discussion to reach consensus or, when necessary, by consultation with a third reviewer

2.8. Certainty of Evidence Assessment

The certainty of evidence for each outcome was assessed using the GRADE approach. Evidence from randomized controlled trials was initially rated as high certainty and could be downgraded based on five domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias. The overall certainty of evidence was classified as high, moderate, low, or very low. GRADE assessments were performed separately for each comparison and treatment duration.

2.9. Statistical Analysis

Meta-analyses were performed using R/RStudio with the *meta* package under a random-effects model. For continuous outcomes, standardized mean differences (SMDs) with Hedges' g correction were used for FSH, E2, AMH, and LH, whereas mean differences (MDs) were used for AFC. For dichotomous outcomes, odds ratios (ORs) were used for clinical effectiveness and risk ratios (RRs) for adverse events. All effect estimates were reported with 95% confidence intervals (CIs), and $P < 0.05$ was considered statistically significant. Statistical heterogeneity was assessed using Cochran's Q test, I^2, and τ^2.

Analyses were stratified by intervention type (TCM vs. HRT; TCM + HRT vs. HRT) and assessment time (≤3 months and >3 months). The 3-month cut-off was selected based on the distribution of intervention durations across the included studies, most of which used treatment courses of 1–3 months or approximately 6 months, and was also consistent with the time-based subgrouping used in previous meta-analyses of traditional Chinese herbal medicine for POI. This stratification was intended to distinguish short-term effects from those observed with longer treatment courses and to explore a potential source of clinical heterogeneity.

Sensitivity analyses were performed where appropriate, including the addition of studies for which means and standard deviations were estimated from medians and interquartile ranges. For meta-analyses with substantial heterogeneity and a sufficient number of studies, leave-one-out analyses were also conducted by sequentially excluding each study to assess its

influence on the pooled effect estimate and I^2. Publication bias was assessed when a sufficient number of studies were available using funnel plots and appropriate tests for asymmetry, including Egger's test for continuous outcomes and Harbord's test for dichotomous outcomes.

3. RESULTS

3.1. Study Selection Process

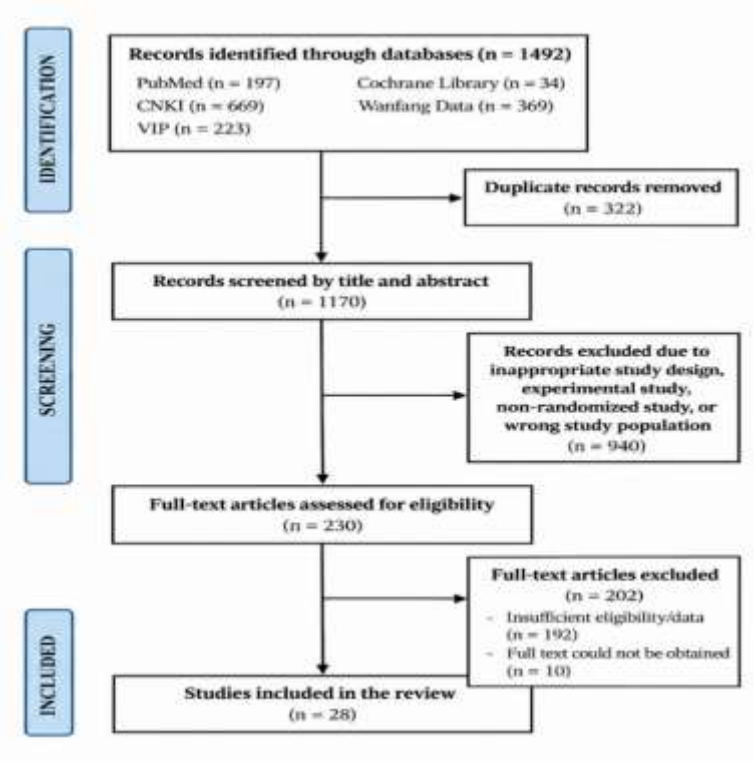


Figure 1.PRISMA Flow Diagram of the Study Selection Process

A total of 1,492 records were identified through the database searches. After screening and eligibility assessment, 28 studies were ultimately included in the systematic review and meta-analysis

3.2. Study Characteristics

Table 1. Baseline Characteristics of the Included Studies

No	Study	Initial sample size	Age (years)	Disease duration	TCM herbal formula	Treatment duration (months)	Outcomes included in this review
1	Tang, 2025 (12)	TH: 30 H: 30	TH: 30,93 ± 3,69 H: 30,23 ± 4,04	TH: 19,83 ± 9,93 H: 18,90 ± 9,65 (months)	Tiaozhou Ziyin Fang	3	1,2,4,5,6,7
2	Xu, 2022a (13)	TH: 40 H: 40	TH: 35,81 ± 2,52 H: 35,81 ± 2,52	TH: 4,16 ± 0,42 H: 4,12 ± 0,63 (years)	Buyin Yishen Tang	2	1,2,4,5,6
3	Chen, 2019 (14)	T: 60 TH: 60 H: 60	NR	NR	Buyin Yangluan Tang	6	1,2,4,5,6,7

4	Ma, 2019 (15)	T: 40 TH: 40 H: 40	T: 29,05 ± 5,37 TH: 29 ± 5,64 H: 29,78 ± 5,41	T: 7,63 ± 2,52 TH: 8,26 ± 3,83 H: 7,97 ± 2,86 (months)	Bushen Shugan	1	1,2,3,5,6
5	Tang, 2021 (16)	T: 40 TH: 40 H: 40	T: 31,63 ± 3,46 TH: 32,93 ± 3,68 H: 32,57 ± 3,60	T: 1,61 ± 0,33 TH: 1,55 ± 0,31 H: 1,58 ± 0,35 (years)	Bushen Huoxue Tang	3	1,2,4,5,6,7
6	Dong, 2024 (17)	T: 30 TH: 30 H: 30	T: 31,26 ± 3,48 TH: 32,19 ± 3,79 H: 31,57 ± 3,66	T: 1,12 ± 0,27 TH: 1,29 ± 0,45 H: 1,23 ± 0,33 (years)	Bushen Huoxue Fang	3	1,2,3,5,6,7
7	Dong, 2020 (18)	T: 41 H: 41	T: 30,97 ± 4,49 H: 30,97 ± 5,83	T: 3,16 ± 1,59 H: 2,95 ± 1,73 (years)	Bushen Huoxue Fang	6	1,2,4,6
8	Wang, 2024 (19)	T: 61 H: 62	T: 32,26 ± 0,54 H: 34,45 ± 0,50	T: 3,75 ± 1,82 H: 3,56 ± 2,07 (years)	Bushen Huoxue Fang	6	1,2,4,6,7
9	Sun, 2025 (20)	TH: 43 H: 43	TH: 33,26±3,45 H: 33,11±3,06	TH: 12,25±2,16 H: 12,31±2,20 (months)	Bugan Yishen Tiaochong	3	1,2,3,4,5,6,7
10	Hao, 2022 (21)	TH: 60 H: 60	TH: 35,42 ± 2,09 H: 35,54 ± 2,12	TH: 2,11 ± 0,46 H: 2,10 ± 0,49 (years)	Buxian Bugan Yangxue	1	1,2,3,4,5,6,7
11	Zhao. 2025 (22)	TH: 40 H: 40	TH: 33,73 ± 2,79 H: 33,72 ± 3,61	TH: 2,49 ± 0,56 H: 2,38 ± 0,72 (years)	Self-formulated Wenjing Tang	3	1,2,3,4,5,7
12	Yin, 2022 (23)	TH: 30 H: 30	TH: 34,87 ± 4,03 H: 34,93 ± 4,35	NR	Yiwei Tang	3	1,2,4,5
13	Zhang, 2022 (24)	TH: 42 H: 42	TH: 33,05±2,94 H: 32,15±3,26	TH: 24,05±6,90 H: 22,71±7,84 (months)	Yishen Fuchao Tang	3	1,2,4,5,6
14	Xu, 2022b (25)	TH: 95 H: 95	TH: 34,8±3,4 H: 35,2±3,6	TH: 35,2±3,6 H: 9,3±3,2 (months)	Shugan Bushen Huoxue	3	1,2,4,5,6
15	Chen, 2025 (26)	TH: 43 H: 43	TH: 34,88 ± 2,95 H: 34,15 ± 3,24	TH: 1,73 ± 0,42 H: 1,97 ± 0,36 (years)	Zishen Shugan Fang	1,5	1,2,3,5,6,7
16	Feng, 2024 (27)	T: 43 H: 43	T: 34,09 ± 3,54 H: 33,35 ± 3,17	T: 13,52 ± 6,17 H: 13,04 ± 5,26 (months)	Jiajian Wenjing Tang	3	1,2,4,5,6
17	Guo, 2023(28)	TH: 46 H: 46	TH: 34,33 ± 4,57 H: 34,22 ± 4,26	TH: 12,57 ± 2,44 H: 12,44 ± 2,11 (months)	Yulin Yanzong Tang	3	1,2,3,5,6,7
18	Lin, 2024 (29)	T: 40 H: 40	T: 32,12 ± 4,07 H: 32,27 ± 4,16	NR	Buxue Peiyuan Tang	3	1,2,4,6,7
19	Song, 2019 (30)	T: 20 H: 20	NR	NR	Combined Chinese	6	1,2,3,4,5,6

					Herbal Formula		
20	Jin, 2013 (31)	T: 86 TH: 91 H: 88	T: 33,25 ± 3,11 TH: 33,32 ± 3,22 H: 34,11 ± 3,31	NR	Bushen Huoxue Fang	6	1,2,3,5,6,7
21	Sui, 2021 (32)	T: 60 H; 62	T: 32,45 ± 0,93 H: 34,61 ± 0,61	T: 2,83 ± 1,36 H: 2,79 ± 1,31	Zhu Shi Tiaojing Fang	6	1,2,4,5,6
22	Li, 2024 (33)	T: 30 TH: 30 H: 30	T: 33,27 ± 6,63 TH: 34,12 ± 6,78 H; 33,85 ± 6,54	T: 24,84 ± 3,58 TH: 24,26 ± 3,67 H: 24,53 ± 3,82 (months)	Xinjia Guishen Wan	3	1,2,3,5,6
23	Song, 2021 (34)	T: 35 H: 38	T: 35,48 ± 0,93 H: 37,23 ± 0,61	NR	Bushen Huoxue Fang	3	1,2,3,4,6,7
24	Li, 2019 (35)	TH: 44 H: 44	TH: 32,47 ± 4,11 H: 32,45 ± 4,13	TH: 15,21 ± 3,16 H: 15,23 ± 3,14 (months)	Qingxin Zishen Tang	3	1,2,4,5,6,7
25	Chi, 2025 (36)	TH: 48 H: 47	TH: 36,56 ± 2,89 H: 37,25 ± 1,89	TH: 54,23 ± 9,50 H: 55,89 ± 3,57 (days)	Zishen Yugui Tang	3	1,2,3,5,6,7
26	Du, 2023 (37)	TH: 46 H: 47	TH: 29,15 ± 3,48 H: 28,78 ± 3,56	TH: 3,58 ± 0,88 H: 3,66 ± 0,96 (years)	Jiajian Wenjing Tang	3	1,2,3,4,5,6,7
27	Lu, 2022 (38)	T: 60 H: 60	T: 33,52 ± 5,42 H: 33,45 ± 5,15	T: 2,15 ± 0,68 H: 2,19 ± 0,65 (years)	Erxian Tang combined with Yougui Tang	3	1,2,4,5,6,7
28	Huang, 2025 (39)	T: 25 H: 25	T: 30 ± 5 H: 30 ± 4	T: 1,5 ± 0,2 H: 1,4 ± 0,3 (years)	Wubie Huankui Tang	3	1,2,4,5,6

Note: T: Traditional Chinese medicine (TCM); TH: TCM combined with hormone replacement therapy; H: hormone replacement therapy; NR: Not reported; 1: serum follicle-stimulating hormone (FSH) levels; 2: serum Estradiol (E_2) levels; 3: antral follicle count (AFC); 4: AMH; 5: the total effective rate; 6: serum luteinizing hormone (LH) levels; 7: adverse events.

3.3. Risk of Bias

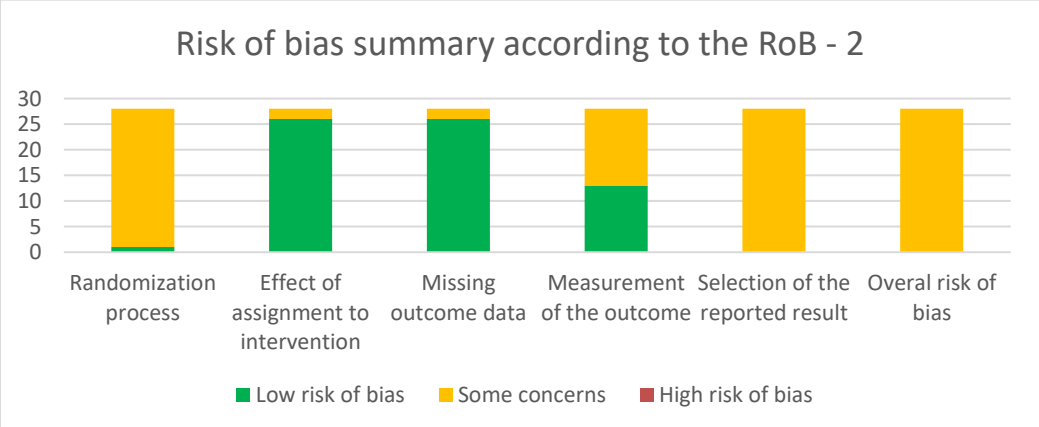


Figure 2. Summary of Risk of Bias Assessments Using the RoB-2 Tool

Among the 28 included studies, 22 described random sequence generation using a random number table, and one study used sealed opaque envelopes for allocation concealment. Two studies reported loss to follow-up or withdrawal. Most studies had complete outcome data and showed no substantial bias due to deviations from the intended interventions. None of the studies was judged to be at high risk of bias; however, all had at least one domain rated as "some concerns," mainly related to the randomization process, outcome measurement, and selection of the reported result.

3.4. Certainty of Evidence Assessment

Table 2. GRADE Assessment of the Certainty of Evidence

Outcome	Comparison / subgroup	Studies (N)	Pooled effect (95% CI)	Certainty of evidence
FSH	TCM- ≤ 3 Month	9 (689)	SMD −0,31 [−0,61; −0,01]	Very low
FSH	TCM, > 3 Month	4 (410)	SMD −0,29 [−1,12; 0,54]	Very low
FSH	TCM + HRT, ≤ 3 Month	17 (1494)	SMD −1,61 [−1,96; −1,26]	Low
FSH	TCM + HRT, >3 Month	3 (334)	SMD −0,69 [−1,36; −0,01]	Very low
E2	TCM- ≤3 Month	9 (689)	SMD 0,44 [−0,26; 1,14]	Very low
E2	TCM, > 3 Month	4 (410)	SMD 0,06 [−1,79; 1,92]	Very low
E2	TCM + HRT, ≤ 3 Month	17 (1494)	SMD 1,07 [0,71; 1,43]	Very low
E2	TCM + HRT, >3 Month	3 (334)	SMD 1,22 [−0,01; 2,45]	Very low
AMH	TCM- ≤ 3 Month	6 (489)	SMD 0,49 [−0,01; 1,00]	Very low
AMH	TCM, > 3 Month	4 (358)	SMD 0,04 [−0,17; 0,25]	Low
AMH	TCM + HRT, ≤ 3 Month	11 (1021)	SMD 1,35 [0,59; 2,11]	Very low
AMH	TCM + HRT, >3 Month	2 (160)	SMD 0,44 [0,13; 0,75]	Low
AFC	TCM- <=3 Month	5 (336)	MD 0,22 [−0,01; 0,45]	Very low
AFC	TCM + HRT, <3 Month	12 (1097)	MD 1,17 [0,78; 1,56]	Very low
Overall clinical effectiveness	TCM- <=3 Month	8 (616)	OR 1,69 [1,05; 2,74]	Very low
Overall clinical effectiveness	TCM, > 3 Month	4 (454)	OR 1,57 [0,82; 3,02]	Very low
Overall clinical effectiveness	TCM + HRT, <=3 Month	16 (1434)	OR 3,56 [2,62; 4,83]	Low
Overall clinical effectiveness	TCM + HRT, >3 Month	3 (339)	OR 2,40 [1,24; 4,65]	Very low
Adverse events	TCM + HRT ≤3 tháng	4	RR 0,73 [0,41; 1,30]	Very low
Adverse events	TCM ≤3 tháng	3	RR 0,92 [0,36; 2,37]	Very low

The certainty of evidence was assessed using the GRADE approach. Because most studies raised some concerns regarding risk of bias, the certainty of evidence was downgraded by one level for this domain. Inconsistency was the most common reason for downgrading, particularly for FSH, E2, AMH, and AFC, for which substantial heterogeneity was observed. Some analyses were further downgraded for imprecision because of the small number of studies or wide confidence intervals. No clear evidence of publication bias was identified in analyses with a sufficient number of studies for assessment. Overall, the certainty of evidence for most outcomes was rated as low or very low.

Natural Resources for Human Health
eISSN: 2583-1194

3.5. Assessment of Treatment Effectiveness

3.5.1. Overall Clinical Effectiveness

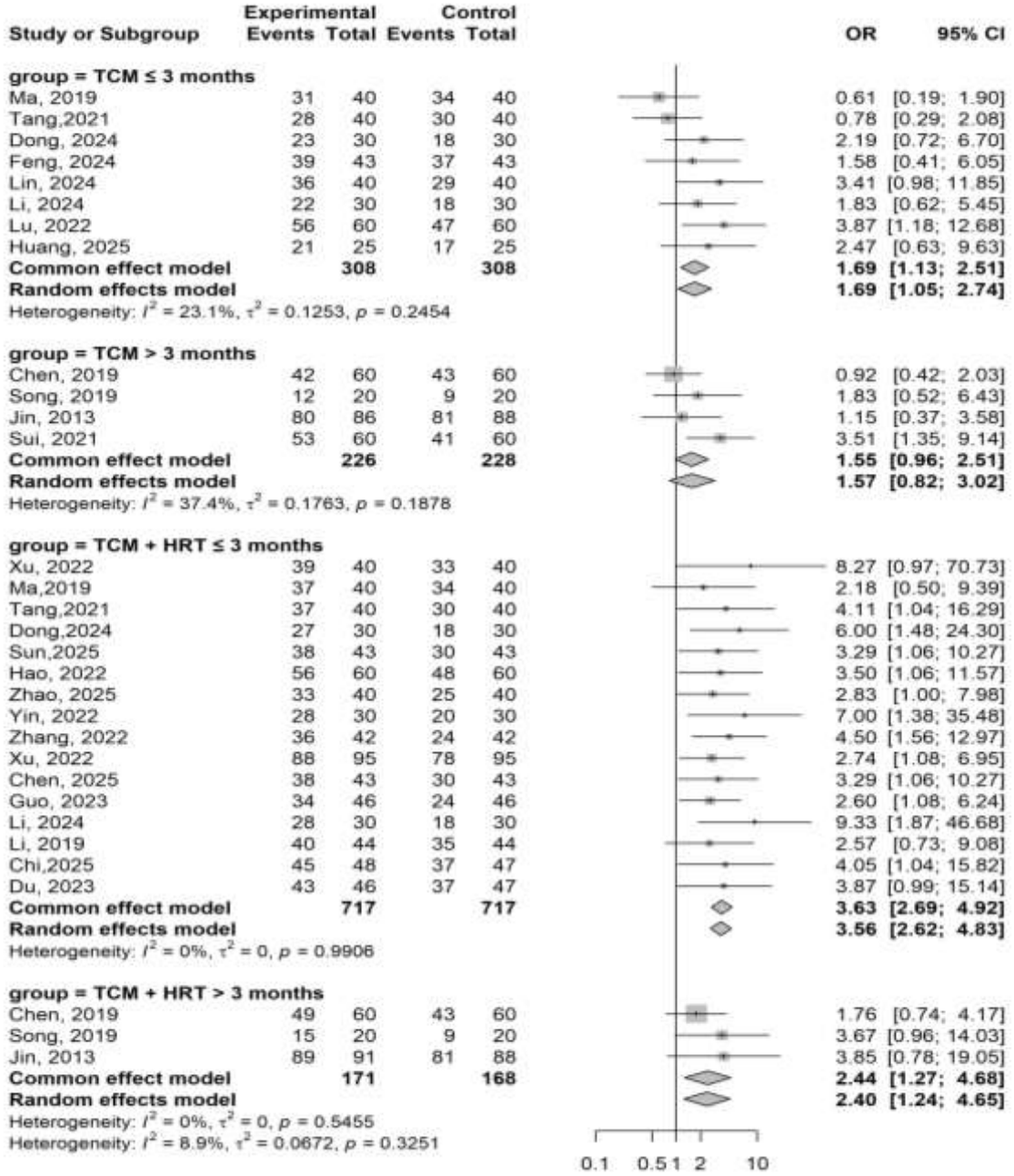


Figure 3. Meta-analysis of Overall Clinical Effectiveness

The meta-analysis showed that TCM combined with HRT significantly improved overall clinical effectiveness compared with HRT alone at both ≤3 months (OR = 3.56; 95% CI: 2.62–4.83; I^2 = 0%) and >3 months (OR = 2.40; 95% CI: 1.24–4.65; I^2 = 0%). For TCM alone, overall clinical effectiveness was also higher than with HRT at ≤3 months (OR = 1.69; 95% CI: 1.05–2.74; I^2 = 23.1%); however, no statistically significant difference was observed at >3 months (OR = 1.57; 95% CI: 0.82–3.02; I^2 = 37.4%).

Natural Resources for Human Health
eISSN: 2583-1194

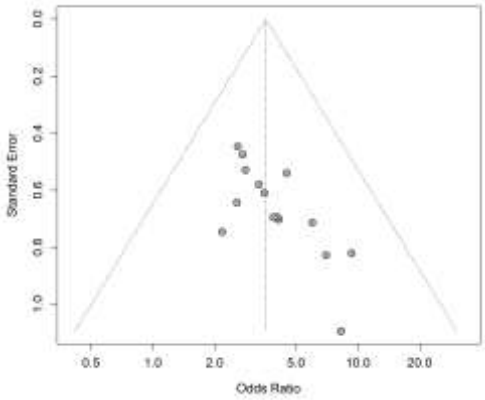


Figure 4. Funnel Plot of Overall Clinical Effectiveness

Harbord's test showed no statistically significant evidence of funnel plot asymmetry for overall clinical effectiveness (t = 1.96; df = 14; P = 0.071).

3.5.2. FSH Levels

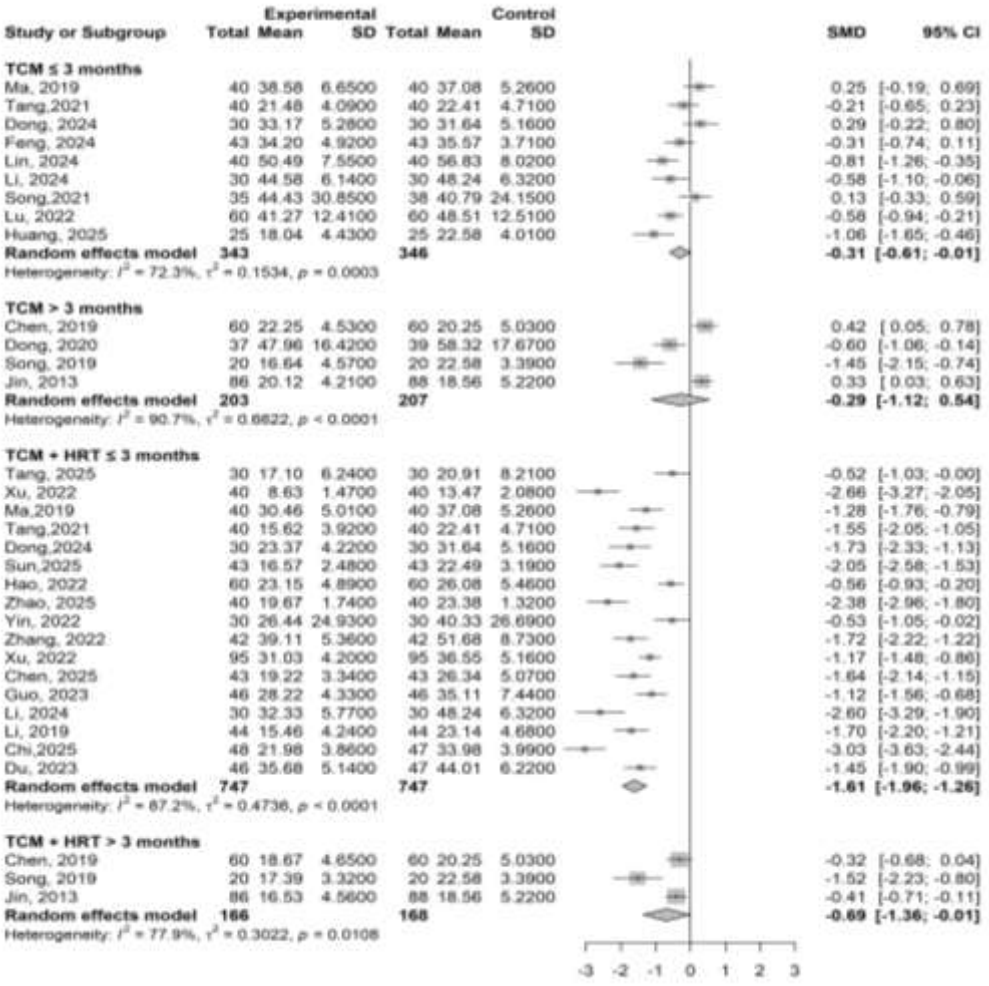


Figure 5. Meta-analysis of FSH Levels

The meta-analysis showed that TCM alone significantly reduced FSH levels compared with HRT in the ≤3-month subgroup (SMD = −0.31; 95% CI: −0.61 to −0.01; I^2 = 72.3%), whereas no statistically significant difference was observed in the >3-month subgroup (SMD = −0.29; 95% CI: −1.12 to 0.54; I^2 = 90.7%).

For TCM combined with HRT, FSH levels were significantly lower than with HRT alone at both ≤3 months (SMD = −1.61; 95% CI: −1.96 to −1.26; I^2 = 87.2%) and >3 months (SMD = −0.69; 95% CI: −1.36 to −0.01; I^2 = 77.9%). Two studies, Wang (2024) and Sui (2021), reported FSH values as medians and interquartile ranges and were therefore excluded from the primary analysis.

When these studies were added in a sensitivity analysis after estimating the corresponding means and standard deviations, the result for TCM alone at >3 months remained statistically non-significant. However, the effect estimate was markedly attenuated from SMD = −0.29 to −0.04 (95% CI: −0.63 to 0.54), indicating that the magnitude of the pooled effect was not fully robust.

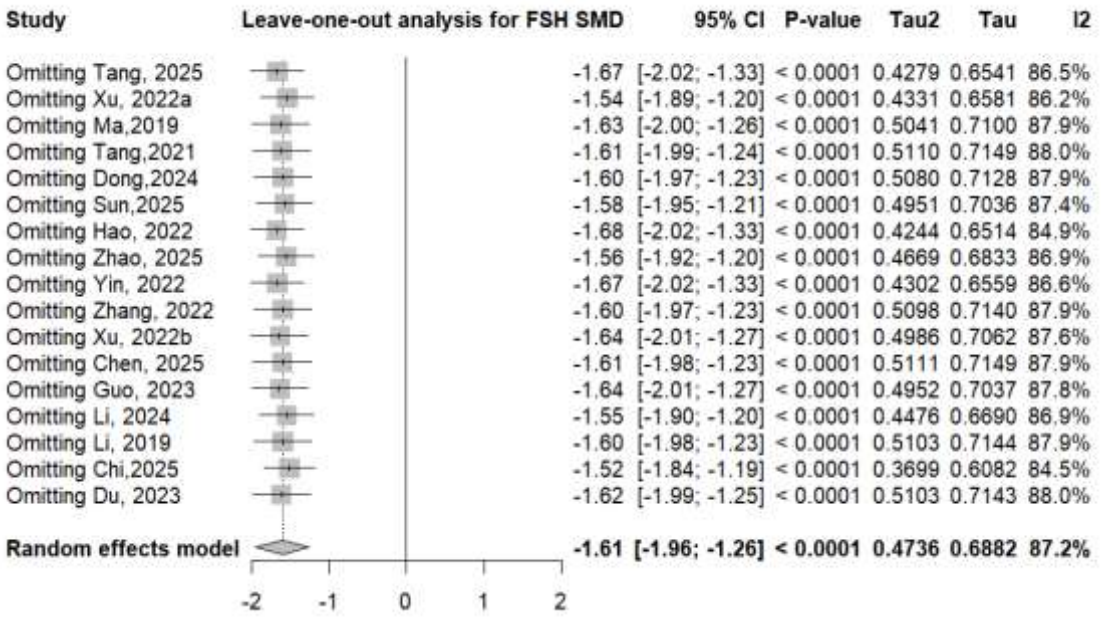


Figure 6. Leave-One-Out Sensitivity Analysis of FSH in the TCM + HRT Group at ≤3 Months

The leave-one-out analysis showed that the pooled result remained stable after sequentially excluding each study (SMDs ranging from −1.68 to −1.52, with all 95% CIs excluding 0). Excluding Chi (2025) produced the largest reduction in heterogeneity, with I^2 decreasing from 87.2% to 84.5%; however, heterogeneity remained high, indicating that no single study was primarily responsible for the observed inconsistency.

3.5.3. E2 Levels

Natural Resources for Human Health
eISSN: 2583-1194

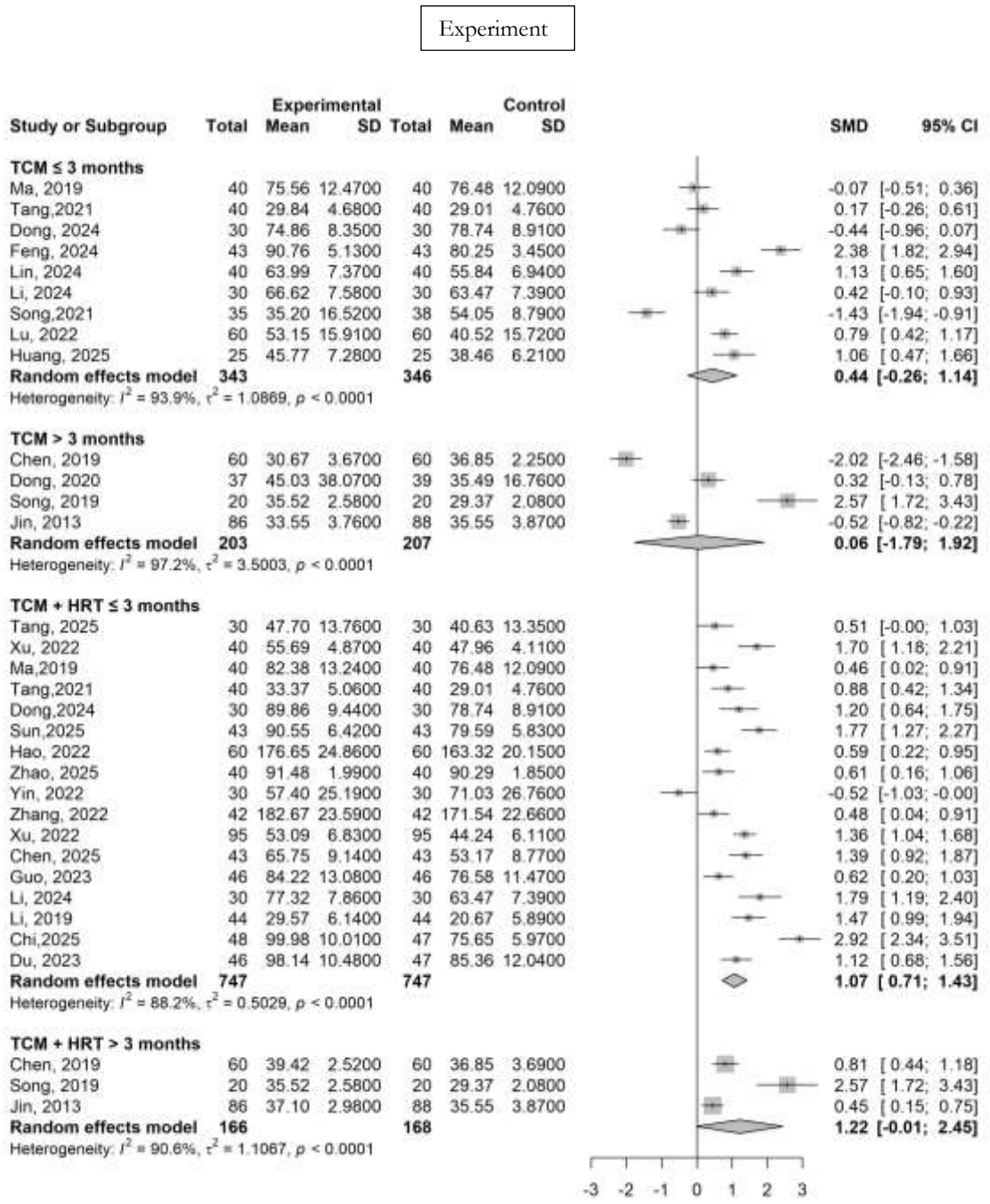


Figure 7. Meta-analysis of E2 Levels

TCM alone did not significantly affect E2 levels compared with HRT in either the ≤3-month subgroup (SMD = 0.44; 95% CI: −0.26 to 1.14) or the >3-month subgroup (SMD = 0.06; 95% CI: −1.79 to 1.92). In contrast, TCM combined with HRT significantly increased E2 levels compared with HRT alone at ≤3 months (SMD = 1.07; 95% CI: 0.71–1.43), whereas no statistically significant difference was observed at >3 months (SMD = 1.22; 95% CI: −0.01 to 2.45). Two studies, Wang (2024) and Sui (2021), reported E2 values as medians and interquartile ranges and were therefore excluded from the primary analysis. Sensitivity analysis including these two studies did not change the conclusion for the TCM-alone group at >3 months (SMD = −0.29; 95% CI: −1.53 to 0.96).

Natural Resources for Human Health
eISSN: 2583-1194

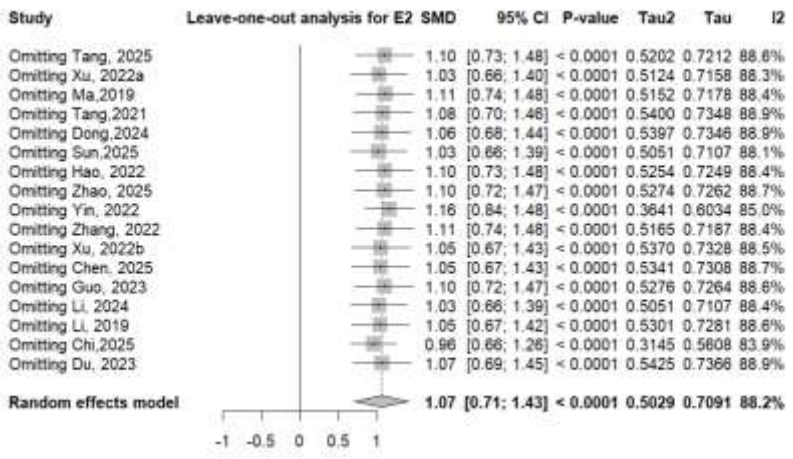


Figure 8. Leave-One-Out Sensitivity Analysis of E2 in the TCM + HRT Group at ≤3 Months

The leave-one-out analysis showed that the pooled estimate remained stable after sequentially excluding each study (SMDs ranging from 0.96 to 1.16, with all 95% CIs excluding 0). Excluding Chi (2025) resulted in the largest reduction in heterogeneity, with I^2 decreasing from 88.2% to 83.9%; however, heterogeneity remained high, indicating that no single study was primarily responsible for the observed inconsistency.

3.5.4. AMH Levels

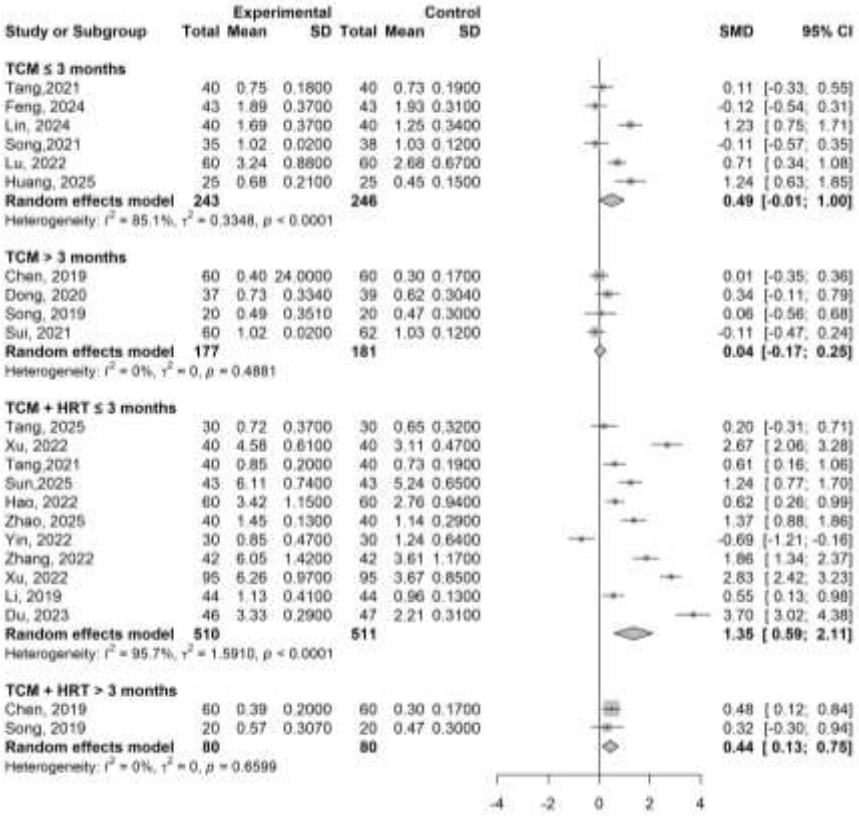


Figure 9. Meta-analysis of AMH Levels

Natural Resources for Human Health
eISSN: 2583-1194

TCM alone did not significantly increase AMH levels compared with HRT in either the ≤3-month subgroup (SMD = 0.49; 95% CI: −0.01 to 1.00) or the >3-month subgroup (SMD = 0.04; 95% CI: −0.17 to 0.25). In contrast, TCM combined with HRT significantly increased AMH levels compared with HRT alone at both ≤3 months (SMD = 1.35; 95% CI: 0.59–2.11) and >3 months (SMD = 0.44; 95% CI: 0.13–0.75). Sui (2021) reported AMH values as medians and interquartile ranges; AMH increased significantly from baseline in both groups (P < 0.001), while the post-treatment median value was descriptively higher in the intervention group than in the control group.

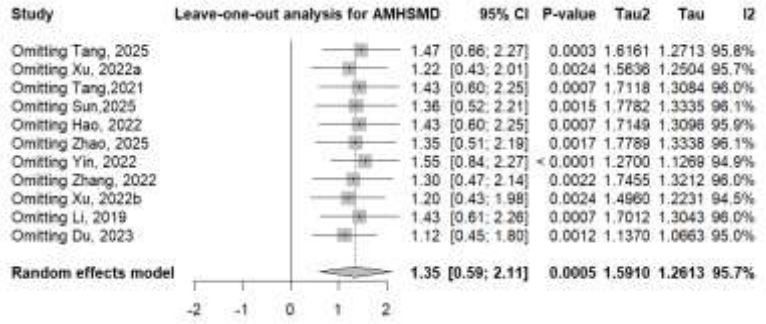


Figure 10. Leave-One-Out Sensitivity Analysis of AMH in the TCM + HRT Group at ≤3 Months

The leave-one-out analysis showed that the pooled estimate remained relatively stable after sequentially excluding each study (SMDs ranging from 1.12 to 1.55, with all 95% CIs excluding 0). I^2 remained very high, ranging from 94.5% to 96.1%, indicating that no single study was primarily responsible for the observed heterogeneity.

3.5.5. LH Levels

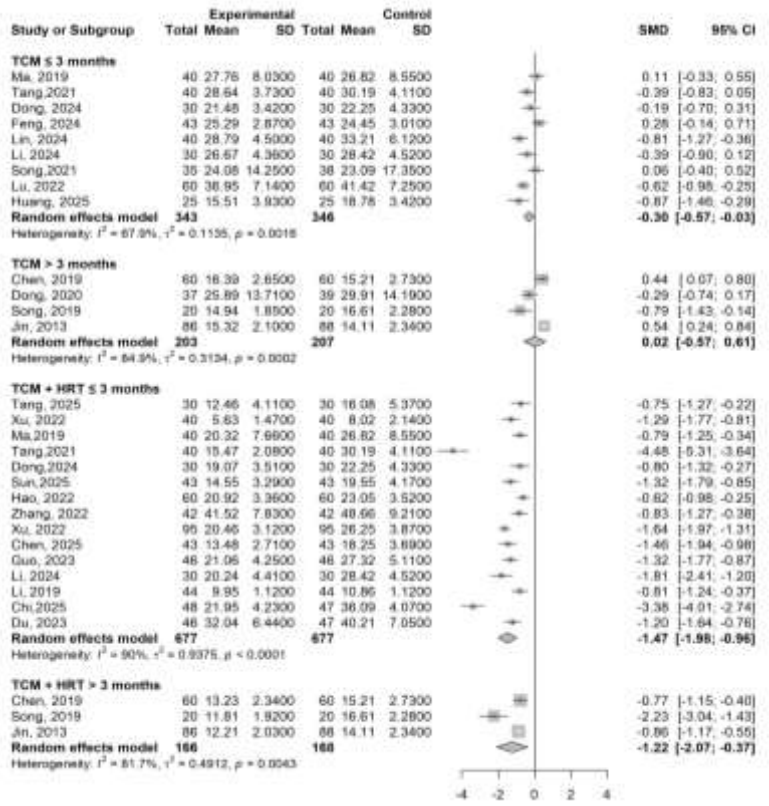


Figure 11. Meta-analysis of LH Levels

TCM alone did not significantly affect LH levels compared with HRT at either ≤3 months (SMD = −0.38; 95% CI: −0.77 to 0.01) or >3 months (SMD = −0.18; 95% CI: −0.76 to 0.40). In contrast, TCM combined with HRT significantly reduced LH levels compared with HRT alone at both ≤3 months (SMD = −1.64; 95% CI: −2.18 to −1.10) and >3 months (SMD = −1.42; 95% CI: −2.24 to −0.61). Sensitivity analysis including Wang (2024) and Sui (2021) did not change the conclusion for the TCM-alone group at >3 months (SMD = 0.07; 95% CI: −0.28 to 0.42), while heterogeneity decreased to I^2 = 75.6%.

3.5.6. AFC

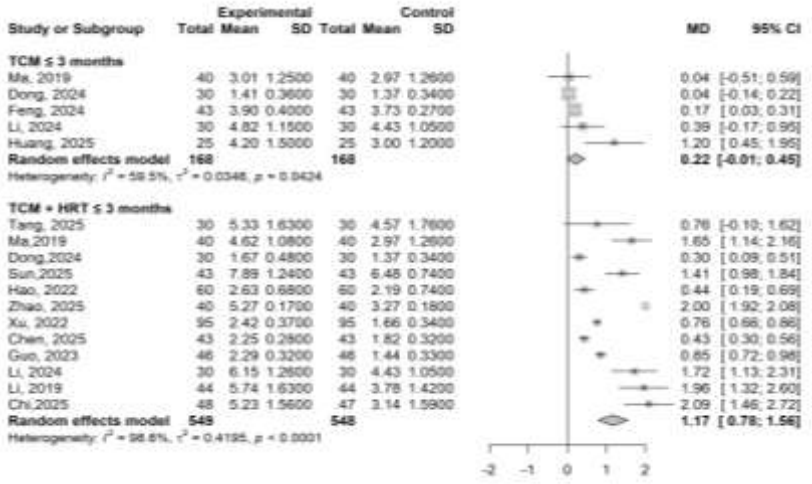


Figure 12. Meta-analysis of AFC

TCM alone did not significantly increase AFC compared with HRT in the ≤3-month subgroup (MD = 0.22; 95% CI: −0.01 to 0.45; I^2 = 59.5%). In contrast, TCM combined with HRT significantly increased AFC compared with HRT alone (MD = 1.17; 95% CI: 0.78–1.56), although substantial heterogeneity was observed across studies (I^2 = 98.6%).

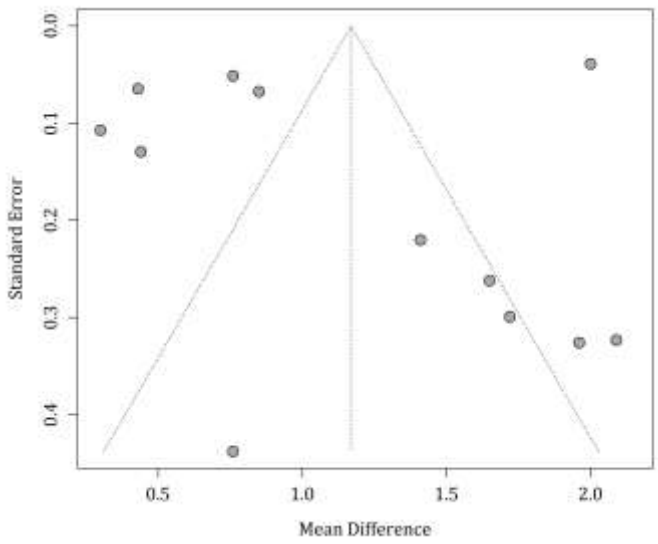


Figure 13. Funnel Plot of AFC

Egger's test showed no evidence of funnel plot asymmetry for the AFC outcome (t = −0.60; df = 10; P = 0.563).

Natural Resources for Human Health
eISSN: 2583-1194

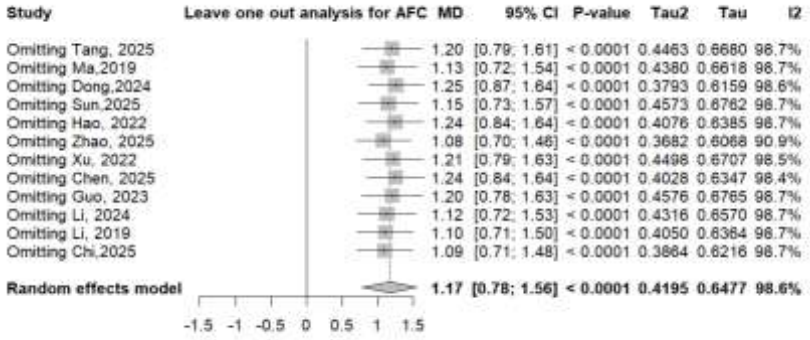


Figure 14. Leave-One-Out Sensitivity Analysis of AFC

The leave-one-out analysis showed that the pooled effect remained relatively stable after sequentially excluding each study (MDs ranging from 1.08 to 1.25, with all 95% CIs excluding 0). Excluding Zhao (2025) resulted in the largest reduction in heterogeneity, with I^2 decreasing from 98.6% to 90.9%; however, heterogeneity remained very high, indicating that no single study was primarily responsible for the observed inconsistency..

3.5.7. Adverse Events

Among the 28 studies included in the review, 17 reported information related to safety or adverse events; however, reporting methods varied across studies. For the comparison of TCM combined with HRT versus HRT alone at ≤3 months, four studies—Tang (2025), Li (2019), Chi (2025), and Du (2023)—provided sufficient data on the total number of patients experiencing adverse events for meta-analysis. No statistically significant difference in the risk of adverse events was observed between the two groups (RR = 0.73; 95% CI: 0.41–1.30; P = 0.291; I^2 = 33.9%). For the comparison of TCM alone versus HRT at ≤3 months, three studies—Lin (2024), Song (2021), and Lu (2022)—were included in the meta-analysis. Similarly, no statistically significant difference was observed between the two groups (RR = 0.92; 95% CI: 0.36–2.37; P = 0.865; I^2 = 9.4%). For treatment durations >3 months, quantitative data were limited. In the comparison of TCM alone versus HRT, Wang (2024) reported adverse events in 2/61 patients in the TCM group and 1/62 patients in the HRT group, whereas Jin (2013) reported no obvious adverse reactions in the treatment groups. For TCM combined with HRT versus HRT alone, Chen (2019) reported individual types of adverse events separately, while Jin (2013) reported no obvious adverse reactions; therefore, no meta-analysis was performed. Several other studies only described specific adverse events, reported safety laboratory findings, or stated that no obvious adverse reactions occurred without providing the total number of patients experiencing adverse events. These studies were summarized descriptively and were not included in the quantitative analysis.

4. DISCUSSION

4.1. Characteristics of the Included Studies

This review included 28 randomized clinical trials evaluating two main treatment strategies: oral TCM alone versus HRT and TCM combined with HRT versus HRT alone. The TCM herbal formulations varied considerably in composition, therapeutic principles, and treatment duration, which generally ranged from 1 to 6 months. Reported outcomes mainly included FSH, E2, AMH, AFC, LH, overall clinical effectiveness, and adverse events. Most studies had relatively small sample sizes and methodological limitations. According to the RoB 2 assessment, all studies had at least one domain rated as "some concerns," mainly related to the randomization process, outcome measurement, and selection of the reported result.

4.2. Main Findings

The findings of this meta-analysis suggest that the effects of TCM varied according to the treatment strategy. TCM alone showed clearer benefits for only some outcomes, mainly within the ≤3-month treatment period, whereas several outcomes, including E2, AMH, LH, and AFC, did not differ significantly from HRT. In contrast, TCM combined with HRT showed more consistent improvements across multiple outcomes, including reductions in FSH and LH, increases in AMH, and improvements

in E2 and AFC in some subgroups, together with a higher overall clinical effectiveness rate. These findings are generally consistent with previous meta-analyses of TCM combined with hormone therapy, in which combination treatment showed benefits for several endocrine outcomes compared with hormone therapy alone (40). A recent meta-analysis comparing oral TCM formulations alone with HRT also reported favorable trends in FSH, AMH, LH, E2, and AFC; however, substantial heterogeneity was observed and the certainty of the evidence was low. (7) Although changes in FSH, E2, AMH, AFC, and LH may reflect improvements in ovarian endocrine function and ovarian reserve, these outcomes are primarily biological markers and do not directly indicate restoration of fertility. Therefore, the present findings should be interpreted mainly as improvements in ovarian function-related parameters rather than evidence of reproductive recovery. The GRADE assessment also indicated that the certainty of evidence was low or very low for most outcomes, mainly because of risk of bias, inconsistency, and imprecision. High heterogeneity was observed in several analyses, particularly for FSH, E2, AMH, and AFC, with I^2 exceeding 90% in some comparisons. This heterogeneity may reflect multiple sources. First, the TCM formulations varied substantially in their ingredients, dosages, and therapeutic principles, which may have resulted in differences in pharmacological effects. Second, the HRT regimens used in the control groups were not fully consistent in terms of hormone type, dosage, and administration. Third, treatment duration ranged from approximately 1 to 6 months, while baseline patient characteristics, including age, disease duration, and severity of ovarian dysfunction, also differed across studies. In addition, the diagnostic criteria for POI/POF varied according to the time of publication, laboratory methods for hormone measurements and AFC assessment may have differed across centers, and the methodological quality of the included RCTs was generally limited. Leave-one-out analyses of comparisons with sufficient numbers of studies and substantial heterogeneity showed that the pooled effects for FSH, E2, AMH, and AFC in the TCM + HRT ≤3-month subgroup remained in the same direction after sequentially excluding each study. However, I^2 generally remained high, suggesting that the observed heterogeneity was not primarily driven by any single study but more likely reflected clinical and methodological differences across trials. Therefore, the pooled estimates should be interpreted as average effects across heterogeneous populations and treatment regimens rather than as uniform effects applicable to all clinical settings. Regarding safety, 17 of the 28 included studies reported information on adverse events, although reporting methods were inconsistent across studies. The available analyses showed no statistically significant differences in the risk of adverse events among TCM, TCM + HRT, and HRT in subgroups with sufficient data. Therefore, the current evidence does not support the conclusion that TCM is safer than HRT; rather, it indicates that no clear difference in adverse event risk has been identified.

4.3. Limitations

This study has several limitations. First, the methodological quality of the included RCTs was generally suboptimal, and all studies had at least one risk-of-bias domain rated as "some concerns." Second, the TCM formulations varied substantially in composition, dosage, and treatment duration, which may have contributed to the high heterogeneity observed in several analyses. Third, some subgroups included only a small number of studies, reducing the precision of the effect estimates and limiting the assessment of publication bias. Fourth, most outcomes were biological markers, whereas patient-important outcomes such as menstrual recovery, pregnancy, live birth, and quality of life were not adequately evaluated. Requiring eligible studies to report both FSH and E2 together with at least one ovarian reserve marker (AMH or AFC) may have excluded some RCTs that otherwise met the PICO criteria but did not measure or fully report these outcomes. Therefore, outcome-related study selection bias cannot be completely excluded. In addition, this review included only full-text articles published in English or Chinese and did not include grey literature; therefore, the possibility of language bias and publication bias cannot be completely ruled out. Finally, adverse event reporting was inconsistent across studies, and follow-up durations were generally short, limiting the ability to fully assess long-term safety.

5. CONCLUSION

This systematic review and meta-analysis suggests that oral traditional Chinese herbal medicine, particularly when combined with HRT, may improve several biological markers of ovarian endocrine function and ovarian reserve, including FSH, E2, AMH, AFC, and LH in some subgroups. However, these changes should not be interpreted as evidence of restored fertility because data on ovulation, pregnancy, and live birth were insufficient. Regarding safety, no clear differences in the risk of adverse events were identified between treatment groups. However, the available evidence remains limited because relatively few studies reported safety outcomes, the number of events was small, and follow-up periods were short. Given the substantial

heterogeneity and the predominantly low or very low certainty of evidence, further well-designed RCTs with larger sample sizes and longer follow-up are needed to confirm the clinical efficacy and safety of traditional Chinese herbal medicine for POI.

6. ACKNOWLEDGMENTS

We sincerely thank the investigators of the trials included in this review for publishing their study findings, which made it possible to synthesize and evaluate the available evidence in the present study.

7. CONFLICT OF INTEREST

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

REFERENCES

[1] Webber L, Davies M, Anderson R, Bartlett J, Braat D, Cartwright B, et.al. ESHRE Guideline: management of women with premature ovarian insufficiency. Human Reproduction. 2016;31(5):926-37.

[2] Li M, Zhu Y, Wei J, Chen L, Chen S, Lai D. The global prevalence of premature ovarian insufficiency: a systematic review and meta-analysis. Climacteric. 2023;26(2):95-102.

[3] 崔 楠 杨琪, 杨亚楠, 王晓甜, 王海霞, 邓楹君, 黄欲晓. 卵巢早衰的中医病机及其方剂现代药理探索. 世界科学技术-中医药现代化★中药学研究. 2022;24(5).

[4] Chon SJ, Umair Z, Yoon MS. Premature Ovarian Insufficiency: Past, Present, and Future. Front Cell Dev Biol. 2021;9:672890.

[5] Armeni E, Paschou SA, Goulis DG, Lambrinoudaki I. Hormone therapy regimens for managing the menopause and premature ovarian insufficiency. Best Pract Res Clin Endocrinol Metab. 2021;35(6):101561.

[6] Guo Q, Ma W-r, Tan Y. Treatment of Premature Ovarian Insufficiency from the Perspective of "the Two-Yang Disease Developing in the Heart and Spleen" Theory. Journal of Nanjing University of traditional Chinese Medicine. 2022;38(8):735-8.

[7] Wu Y LS, Liu G, Zheng X, Cheng J, Jiang X Oral Chinese herbal formulas for premature ovarian insufficiency: a systematic review, meta-analysis, and exploratory therapeutic-principle subgroup analysis. Front Endocrinol. 2026;17.

[8] Zhong HZ, Li MY, Yin XL, Bin CL, Zhou SY, Wei SB. Chinese patent medicines combined with hormone replacement therapy for premature ovarian failure: A Bayesian network meta-analysis. Front Med (Lausanne). 2022;9:1043390.

[9] Lin F, Lian J, Wu L. Examining the effect of Kuntai capsule combined with HRT in the treatment of premature ovarian insufficiency: A systematic review and meta-analysis. Exp Ther Med. 2025;29(6):116.

[10] Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Bmj. 2021;372:n71.

[11] Nejadghaderi SA, Balibegloo M, Rezaei N. The Cochrane risk of bias assessment tool 2 (RoB 2) versus the original RoB: A perspective on the pros and cons. Health Sci Rep. 2024;7(6):e2165.

[12] 唐培培; 谈勇; 殷燕云;黄菁宇;聂晓伟;左文婷; et.al. 调周滋阴方治疗早发性卵巢功能不全的疗效、安全性及作用机制. 南方医科大学学报. 2025;45(5).

[13] 徐彭丽. 补阴益肾汤加减联合雌孕激素序贯疗法治疗卵巢早衰的效果及对 INHB、β-EP 水平的影响. 临床医学研究与实践. 2022;7(6):138-41.

[14] 陈凤萍; 陈耿祝. 补肾调经方联合激素替代治疗卵巢早衰的临床研究. 中国妇幼保健. 2019;34(8):1824-7.

[15] 马平兰; 林寒梅;逯克娜; 贺恒祯; 梁素梅. 补肾疏肝方联合激素替代疗法治疗肾虚肝郁型卵巢早衰 40 例. 河南中医. 2019;39(4):603-7.

[16] 汤海霞 周, 陈晓玲,洪宏. 补肾活血汤联合芬吗通治疗早发性卵巢功能不全(肾虚血瘀型)的疗效及作用机制. 中药材. 2021;44(12):2955-9.

[17] 董亚兰; 张晟;余健秋;王瑞敏. 补肾活血方联合雌二醇片/雌二醇地屈孕酮片对卵巢早衰患者中医证候、卵巢储备功能及 T 细胞亚群的影响. 中国合理用药探索. 2024;21(7):1-7.

[18] 董亚兰; 董莉. 补肾活血方治疗卵巢早衰的临床观察及其对 ANA-ACA-AOA 通路的影响. 上海中医药杂志. 2020;54:61-5.

[19] 王怡青; 黄宏丽;夏艳秋;眭瑾; 董莉. 补肾活血方对肾虚血瘀型早发性卵巢功能不全患者免疫平衡的调节作用. 天津中医药. 2024;41(5):559-64.

[20] 倪芳英; 孙王. 补肝益肾调冲汤治疗早发性卵巢功能不全患者的疗效及对 AMH 和 FSH 水平的影响. 中国性科学. 2025;34(8):108-13.

[21] 郝秀珍; 吴作君;高琦;张素勤. 自拟补仙补肝养血养荣汤治疗 120 例女性卵巢早衰的临床观察. 基层中医药. 2022;1(4):35-41.

[22] 赵俊慧; 秦蕾;范泽玲;张燕南;王玲玲. 自拟温经汤联合芬吗通治疗肾阳虚型卵巢早衰临床观察. 湖北中医药大学学报. 2025;27(6).

[23] 殷增兰;张英杰. 益胃汤治疗早发性卵巢功能不全疗效观察. 实用中医内科杂志. 2022;36(2):67-9.

[24] 张锦聪; 潘蕾;郭凤琴. 益肾复巢汤治疗卵巢早衰/卵巢功能不全患者的疗效及对血清性激素水平的影响. 检验医学与临床. 2022;19(3).

[25] 徐晶;李莹. 疏肝补肾活血方联合激素替代治疗卵巢早衰的效果及其对卵巢储备功能和骨代谢的改善作用研究. 中国性科学. 2022;31(10).

[26] 陈英华. 滋肾疏肝方联合雌孕激素序贯疗法治疗肾虚肝郁型卵巢早衰临床研究. 新中医. 2025;57(3).

[27] 冯艳奇; 马素娟. 温经汤加减治疗卵巢早衰临床研究. 河南中医. 2024;44(9):1316-20.

[28] 郭彩玲 李, 李兴华. 毓麟衍宗汤联合雌孕激素序贯疗法治疗肝郁肾虚型卵巢早衰临床研究. 新中医. 2023;55(22):105-10.

[29] 林娜. 中药序贯治疗卵巢早衰 80 例临床研究. 福建医药杂志. 2024;46(8).

[30] 宋清霞; 邵雪婷;杨晓乐;孙淑梅. 中药复方调整阴阳平衡治疗卵巢早衰的疗效观察. 浙江临床医学. 2019;21(1):56-7.

[31] 金志春 黄, 杨雅琴,王璐,何丹娟,刘文静. 补肾活血方联合雌孕激素治疗 卵巢早衰的临床研究. 中国中西医结合杂志. 2013;33(5).

[32] 眭瑾;董莉. 朱氏调经方对肾虚血瘀型 POI 患者卵巢血供的改善作用. 上海中医药杂志. 2021;55(11).

[33] 李晓红;李芳;龙旭;杨卓. 新加归肾丸联合雌二醇片/雌二醇地屈孕酮片对卵巢早衰患者卵巢功能、Kupperman 评分的影响. 中国性科学. 2024;33(2).

[34] 宋靖宜;夏艳秋;董莉. 基于 AMH 与 microRNA-23a 相关性探讨朱氏补肾活血方对卵巢早衰的影响. 上海中医药杂志. 2021;55(6).

[35] 李娟;张燕;丁志云. 基于"心-肾-子宫轴"理论探讨清心滋肾汤对心肾不交型早发性卵巢功能不全的疗效及部分作用机制. 世界中医药. 2019;14(9).

[36] 池雷;周云. 基于"乙癸同源"理论探究滋肾育癸汤治疗卵巢早衰患者的疗效观察. 世界中西医结合杂志. 2025;20(1):142-6.

[37] 杜嫦燕;关凤仪;罗宝玲. 加减温经汤对脾肾阳虚型卵巢早衰相关激素水平及预后的影响. 中华中医药学刊. 2023;41(8).

[38] 芦希艳;史恒军;姚家久. 二仙汤合右归汤联合西药治疗卵巢早衰患者的临床疗效. 医学临床研究. 2022;39(9).

[39] 黄风; 郭齐;陶文华;王彰禹. 乌鳖还闺汤治疗脾肾亏虚型早发性卵巢功能不全 25 例. 湖南中医杂志. 2025;41(10).

[40] Singh. R., Verghese. S. P., Bansal A. (2026). Silver Nanoparticles For Environmental Sustainability And Their Role In Biological Activities. International Journal of Engineering Sciences & Research Technology, 15(3), 1–8. ttps://doi.org/10.64149/j.ijesrt.15.3-1-8

[41] Kou MJ, Ding XF, Chen JX, Liu Y, Liu YY. Traditional chinese medicine combined with hormone therapy to treat premature ovarian failure: a meta-analysis of randomized controlled trials. Afr J Tradit Complement Altern Med. 2016;13(5):160-9.

APPENDIX 1

SEARCH STRATEGY

The literature search was conducted from the inception of each database to 27 July 2026 in PubMed, the Cochrane Library, China National Knowledge Infrastructure (CNKI), Wanfang Data, and the VIP Database. The search strategy combined controlled vocabulary terms, including Medical Subject Headings (MeSH) where applicable, with free-text terms related to premature ovarian insufficiency and traditional Chinese herbal medicine. The Boolean operators AND and OR were used, and the search syntax was adapted to each database. Studies published in English or Chinese were considered.

1. Wanfang: ("Premature Ovarian Insufficiency" OR "Premature Ovarian Failure" OR "Primary Ovarian Insufficiency") AND ("Chinese Herbal Medicine" OR herb* OR decoction OR "Chinese herbal formula")

2. Cochraine Library : ("Premature Ovarian Insufficiency" OR "Premature Ovarian Failure" OR "Primary Ovarian Insufficiency") AND ("Chinese Herbal Medicine" OR herb* OR decoction OR "Chinese herbal formula")

2. Pubmed: ("Primary Ovarian Insufficiency"[MeSH] OR "primary ovarian insufficiency"[Title/Abstract] OR "premature ovarian insufficiency"[Title/Abstract] OR "premature ovarian failure"[Title/Abstract] OR "primary ovarian failure"[Title/Abstract]) AND ("Drugs, Chinese Herbal"[MeSH] OR "Chinese herbal*"[Title/Abstract] OR herb*[Title/Abstract] OR decoction*[Title/Abstract])

3. VIP: M=("早发性卵巢功能不全" OR "卵巢早衰" OR "原发性卵巢功能不全" OR POI OR POF) AND M=("中药" OR "中草药" OR "中药复方" OR "中药方剂" OR "中药汤剂" OR "中药治疗")

4. CNKI: SU=('早发性卵巢功能不全'+'卵巢早衰'+'原发性卵巢功能不全'+POI+POF)

and SU=('中药'+'中草药'+'中药复方'+'中药方剂'+'中药汤剂'+'中药治疗')

Kết quả tìm kiếm

Database	Date searched	Coverage	Results
PubMed	27-Jul-26	inception–27 Jul 2026	197
Cochrane Library	27-Jul-26	inception–27 Jul 2026	34
CNKI	27-Jul-26	inception–27 Jul 2026	669
Wanfang Data	27-Jul-26	inception–27 Jul 2026	369
VIP	27-Jul-26	inception–27 Jul 2026	223

All records were imported into Zotero. Duplicate records were identified using Zotero's duplicate-detection function and subsequently checked manually before title and abstract screening.