World Journal of Gastroenterology, Hepatology and Endoscopy ISSN 2766-788X | Volume 31

Research Article Open Access

Transcutaneous Auricular Vagus Nerve Stimulation for Mild-To-Moderate Active Crohn's Disease: A Sham-Controlled Crossover Trial

Zixuan Zhou[1*], Hao Su[2*], Lingxiao Huang[1], Siqi Wu[1], Aoxun Zhu[3], Su Xu[1#]

[1]Yancheng TCM Hospital Affiliated to Nanjing University of Chinese Medicine, Department of Anorectal Surgery, Jiangsu, China

[2]Yancheng TCM Hospital Affiliated to Nanjing University of Chinese Medicine, Department of Procurement and Supply Center, Jiangsu, China

[3]Yancheng TCM Hospital Affiliated to Nanjing University of Chinese Medicine, Research Center of Chinese Medicine, Jiangsu, China

[*]These authors contributed equally to this work.

***Corresponding author:**

Su Xu,
Yancheng TCM Hospital Affiliated to Nanjing University of Chinese Medicine, Department of Anorectal Surgery, Jiangsu, China

Received: 03 Oct 2026
Accepted: 26 Oct 2026
Published: 29 Oct 2026
J Short Name: WJGHE

Copyright:

©2026 Su Xu. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and build upon your work non-commercially

Keywords:

Crohn's Disease; Transcutaneous Auricular Vagus Nerve Stimulation; Electrical Nerve Stimulation; Remote-Controlled; Randomized Controlled Trial

Citation:

Su Xu, Transcutaneous Auricular Vagus Nerve Stimulation for Mild-To-Moderate Active Crohn's Disease: A Sham-Controlled Crossover Trial.World Jour of Gastro and Hepatology® 2026; V31(8): 1-13

1. Abstract

1.1. Objective

To evaluate the clinical efficacy and safety of transcutaneous auricular vagus nerve stimulation (taVNS) delivered via a remotely controlled vagus nerve stimulator, versus sham stimulation (sham taVNS), in patients with mild-to-moderate active Crohn's disease (CD).

1.2. Methods

This trial was a single-center, randomized, single-blind, sham-controlled clinical study. A total of 40 patients with mild-to-moderate active CD were enrolled and randomly assigned in a 1:1 ratio to groups A and B, which received different treatment sequences. Treatment was administered once daily for 12 weeks, followed by a 12-week follow-up, giving a total study duration of 24 weeks. Using 4-week cycles, group A received taVNS during weeks 1-4, sham taVNS during weeks 5-8, and taVNS during weeks 9-12; group B received sham taVNS during weeks 1-4 and taVNS from week 5 through the end of the 12-week treatment period. Blood and stool samples, questionnaire data, and dynamic electrocardiogram assessments were collected from all participants at the end of weeks 0, 4, 8, 12, and 24. All efficacy outcome measures (primary: CDAI score; secondary: CRP, ESR, FC, IBDQ, HADS, and HRV) were analyzed using both intention-to-treat (ITT) and per-protocol (PP) analyses.

1.3. Results

A total of 40 patients with mild-to-moderate active CD (20 in group A and 20 in group B) were enrolled; 37 (92.5%) completed all five time-point assessments and were included in the PP analysis set (19 in group A and 18 in group B), while the ITT analysis imputed missing data for the 3 dropouts using last-observation-carried-forward (LOCF). The primary efficacy analysis was a within-subject crossover paired comparison (pairing the change during the true-stimulation period with that during the sham-stimulation period, via the Wilcoxon signed-rank test). Compared with the sham period, the true-stimulation period showed significant improvements in the median CDAI difference of -29.4 points (Hodges-Lehmann -25.1, 95% CI -35.9 to -8.6, P=0.004), median CRP difference of -3.5 mg/L (P=0.008), median ESR difference of -17.0 mm/h (P=0.009), median LF/HF difference of -0.60 (P=0.002), and HFnu (P=0.002), all of which remained significant after Holm correction; in contrast, FC, HADS, and IBDQ showed no statistically significant differences. Concurrent between-group comparisons showed no statistically significant between-group differences at any time point for all other measures except HRV LF/HF (weeks 4 and 8, ITT P=0.004) (all P>0.05). Between-group comparisons of changes (Delta = week 4 - baseline) showed that at week 4 the between-group Delta differences for CDAI, CRP, and ESR were all statistically significant (P=0.018, 0.014, and 0.007, respectively); after Quade baseline correction, CRP (P=0.016), ESR (P=0.029), and HF (P<0.001) remained significant, whereas CDAI became borderline (P=0.054). Regarding within-group comparisons versus baseline, improvements in CRP and ESR were the most robust (CRP remained significant after PP-correction at weeks 4, 12, and 24 in group A and weeks 12 and 24 in group B; ESR remained significant after correction at weeks 4, 8, 12, and

24 in group A and weeks 8, 12, and 24 in group B); FC remained significant after correction at week 24 in both groups and at week 12 in group A (PP analysis). No serious adverse events occurred in terms of safety.

1.4. Conclusion

In patients with mild-to-moderate active CD, taVNS significantly reduced disease activity (CDAI) and inflammation levels (CRP, ESR) and shifted the sympathovagal balance toward vagal predominance (decreased LF/HF and increased HFnu) compared with sham stimulation (within-subject crossover paired comparison), with good safety and tolerability.

2. Introduction

Crohn's disease (CD) is a chronic inflammatory gastrointestinal disorder in which inflammation can affect any part of the digestive tract from the mouth to the anus. The main symptoms include diarrhea and abdominal pain, as well as systemic manifestations such as fever, weight loss, and malnutrition, accompanied by characteristic pathological lesions such as intestinal fibrosis or fistulas [1]. In addition, extraintestinal manifestations involving the joints, skin, or eyes are also common [2]. The global incidence of CD continues to rise, particularly in newly industrialized countries including China, which may be related to the effects of Westernized diet and environmental factors on the gut microbiota [3, 4]. Disease progression seriously impairs patients' quality of life and social functioning. Current CD medications include [5] aminosalicylic acid, glucocorticoids, immunosuppressants, and biologics such as TNF-α antibodies and IL-12/23p40 antibodies5. The continuous introduction of new therapeutic agents has expanded treatment options but also imposed a substantial economic burden, and many patients fail to achieve satisfactory outcomes because of inadequate efficacy or adverse drug reactions; approximately 50% of patients still require surgery to alleviate symptoms due to disease progression within 10 years of diagnosis [6]. Therefore, identifying a safe, effective, and low-cost approach to control disease progression-especially for patients with mild disease activity and lower risk-is an urgent priority for researchers.

The vagus nerve-mediated 'inflammatory reflex' is an important physiological mechanism by which the body regulates peripheral inflammation through the nervous system [7]. When vagal afferent fibers sense peripheral inflammatory signals, the central nervous system releases acetylcholine through vagal efferent fibers, which acts on the α7 nicotinic acetylcholine receptor on immune cells to inhibit the release of pro-inflammatory cytokines [8]. Since Borovikova's team first described this cholinergic anti-inflammatory pathway in 20009, numerous studies have confirmed the key role of autonomic and enteric neural signaling pathways in intestinal immune regulation, inflammation control, and intestinal epithelial barrier function [10,11].

Building on these findings, the field of bioelectronic medicine has sought to apply neuromodulation techniques to the treatment of inflammatory diseases. Preliminary studies of implantable vagus nerve stimulation in Crohn's disease have demonstrated favorable efficacy and safety [12]. In an open-label study in adults with CD, Sinniger et al. found that implantable vagus nerve stimulation improved clinical symptoms, inflammatory biomarkers, and endoscopic healing [13]. D'Haens et al. [12] also confirmed the clinical benefit of implantable VNS in a multicenter study of biologic-refractory moderate-to-severe CD.

However, implantable devices have limitations such as surgical risks, high cost, and low patient acceptability. Transcutaneous auricular vagus nerve stimulation (taVNS) noninvasively activates central vagal projections by stimulating the auricular branch of the vagus nerve in the cymba conchae of the external ear, and functional MRI studies have confirmed that it activates the same brain regions as implantable devices [14]. Proof-of-concept studies of taVNS in rheumatoid arthritis and systemic lupus erythematosus have shown anti-inflammatory potential [15]. In a pediatric IBD proof-of-concept trial, Sahn et al. demonstrated that taVNS improved clinical symptoms and reduced fecal calprotectin levels [16]. Although taVNS has made progress in the treatment of CD, large-scale, long-term taVNS trials in adult CD populations have not yet been established, and the factors influencing its clinical efficacy remain incompletely understood.

To further investigate the clinical efficacy of taVNS in patients with mild-to-moderate Crohn's disease, we independently developed a remotely controlled auricular vagus nerve stimulator and designed a clinical trial consisting of a 12-week treatment period plus a 12-week follow-up, in which patients with mild-to-moderate CD were treated with alternating taVNS neurostimulation and sham taVNS. Based on the stimulator's output parameters, the minimum energy output was used as the standard, setting the output current intensity at 0.1 mA as the stimulation dose for the sham taVNS group [17]. This design was used to evaluate the efficacy and safety of taVNS in patients with mild-to-moderate CD and to examine the influence of different taVNS stimulation protocols on its efficacy in this population.

3. Material and Methods

3.1. Study Design and Participant Recruitment

This trial was a randomized, single-blind, sham-controlled clinical study designed by the investigators. Patients were randomly assigned to receive one of two treatment sequences, consisting of taVNS and sham taVNS in different orders. Long-term efficacy was assessed at 24 weeks. The trial was conducted at the single-center Yancheng TCM Hospital Affiliated to Nanjing University of Chinese Medicine in Jiangsu, China, from February 2025 to August 2026. The study protocol was approved by the Institutional Review Board (approval number KY250413) and registered with the Chinese Clinical Trial Registry (ChiCTR identifier: ChiCTR2500105702). All participants provided written informed consent.

3.2. Inclusion Criteria

1. Confirmed diagnosis of CD within the past 2 years (Montreal classification L1 ileal, L2 colonic, or L3 ileocolonic were all eligible) 18.

2. Age 18-60 years.

3. Patients with mild-to-moderate active disease. Definition of mild-to-moderate active CD: Crohn's Disease Activity Index (CDAI) score between 150 and 450; CDAI score 150-220 is defined as mild disease, 220-450 as moderate, and >450 as severe; CDAI score <150 is defined as clinical remission, and a CDAI decrease of >=100 from baseline is defined as clinical response or treatment efficacy19.

4. Before enrollment, patients were required to have used their current medications (including 5-aminosalicylic acid agents and glucocorticoids) continuously for at least 6 weeks, with a stable drug dosage. If a patient was in a dose-reduction or discontinuation phase, the reduction phase had to be stable for at least 2 weeks.

5. Patients had to voluntarily participate in the study, sign the informed consent form, and fully understand the study objectives, procedures, and potential risks and benefits.

3.3. Exclusion Criteria

1. Severe intestinal complications: patients with serious complications such as intestinal perforation or intestinal obstruction (note: a history of prior abdominal surgery was not an exclusion criterion).

2. Presence of severe cardiovascular disease (e.g., heart failure, arrhythmia) or patients with an implanted cardiac pacemaker, which may affect the measurement of heart rate variability (HRV).

3. Neurological disorders or history of epilepsy: patients with central nervous system diseases or a history of epilepsy.

4. Psychiatric disorder history: patients with severe psychiatric disorders (e.g., schizophrenia, major depression, bipolar disorder) that might interfere with outcome assessment.

5. Pregnant or lactating women.

6. Other severe chronic diseases: patients with severe hepatic or renal insufficiency, or other serious chronic diseases that made them unsuitable for the study, such as diabetic complications or chronic obstructive pulmonary disease.

3.4. Withdrawal Criteria

During the study, patients were withdrawn if disease worsening required adjustment of the original medication or treatment regimen, or if the disease progressed to severe active phase (CDAI > 450) requiring a change in therapeutic strategy. Patients could voluntarily withdraw from the study at any time and for any reason.

3.5. Randomization and Blinding

Randomization was performed by an independent statistician using a computer-generated random number table. The randomization assignments were sealed in opaque envelopes and kept and opened by research personnel not involved in outcome assessment.

This study adopted a single-blind design: participants were unaware of the type of intervention they received (taVNS or sham taVNS). Because sham taVNS differed from taVNS only in that its stimulation current intensity was 0.1 mA, with all other parameters identical, participants could not reliably distinguish the two stimulation conditions through subjective sensation. The adequacy of blinding was assessed at the end of the treatment period (by asking participants to guess which intervention they had received). The investigators operating the device knew the current output settings, but the outcome assessors (personnel performing CDAI scoring and HRV measurements) and the statistical analyst remained blinded to group assignment.

3.6. taVNS Device

3.6.1. Hardware System Design

The remotely controlled auricular vagus nerve stimulator used in this study was independently developed, and its hardware architecture is divided into three levels:

Control core: An STM32F103 microcontroller (MCU) serves as the main control chip, responsible for the digital processing of stimulation parameters and overall system coordination. Wireless data communication with the mobile phone APP is achieved through the WH-BLE106 low-power Bluetooth module, supporting command reception and device status feedback.

Signal stage: A built-in 12-bit digital-to-analog converter (DAC) in the MCU generates a high-precision baseline voltage waveform as the reference signal source for the stimulation pulses, and can flexibly output waveform signals of different frequencies and amplitudes according to preset parameters.

Drive Stage: The core is a pair of improved complementary Wilson current-mirror circuits that efficiently and linearly convert the voltage signal output by the DAC into a constant current independent of load impedance variations. This stage is powered by a high-voltage supply to ensure accurate output current under different auricular skin impedance conditions, providing a stable energy basis for stimulation.

Output stage: An H-bridge full-bridge switching circuit modulates the constant-current signal into asymmetric, bidirectional, charge-balanced stimulation pulses under the control of the precise timing signals generated by the MCU. The core safety advantage of this design is that the positive and negative charges are completely neutralized within each stimulation cycle, fundamentally preventing tissue damage and electrode polarization caused by residual charge accumulation, thereby ensuring the safety of long-term treatment.

3.6.2. Software System Design

The companion APP adopts a front-end/back-end separated architecture, designed to provide physicians with remote patient management and treatment control functions while supporting patient self-treatment: the front-end application (based on the uni-app framework) is compatible with Android, HarmonyOS, and iOS platforms, with role-based login/registration for physicians and patients. Main functions of the physician end: (1) Patient management: binding patients, viewing the patient list, and displaying details; (2) Remote control: real-time monitoring of device connection status, remote setting of treatment parameters (frequency, pulse width, current intensity, stimulation/rest time), and remote session management; (3) Treatment records: viewing treatment history and chart-based statistical analysis by day/week/month; (4) Device management: Bluetooth device binding and status mon-

itoring. Main functions of the patient end: (1) Receiving remote stimulation control: the patient-end APP receives stimulation parameters remotely sent by the bound physician (the parameters are transmitted to the stimulator via the patient's mobile Bluetooth to adjust the stimulation current, but are not displayed to the patient); (2) Aborting stimulation: patients can abort stimulation on their own if they feel discomfort or encounter other emergencies. If a patient voluntarily aborts stimulation, the physician end is also notified.

Back-end service (based on the Flask framework): provides RESTful API interfaces to enable complete data interaction; manages the persistent storage of patient information, treatment records, and device status; implements strict permission verification and patient data isolation mechanisms; and maintains full-lifecycle management of remote control sessions. Security design features: patient data isolation ensures that physicians can access only authorized patient information; remote sessions have creation, verification, and cleanup mechanisms; and Bluetooth device binding requires a device authentication process.

3.6.2. taVNS Stimulation Parameters

Electrical stimulation was delivered using a transcutaneous auricular vagus nerve electrical stimulation system and a special ear-clip electrode. Conductive gel was applied to the electrodes. taVNS was applied to the auricular region, targeting the cymba conchae (kidney CO10) and cavity of concha (heart CO15) on the left ear of the subject-regions rich in vagal innervation-for electrical stimulation20. The stimulation system was remotely set and dynamically adjusted by physicians through the APP, delivering continuous asymmetric biphasic square pulse waves. The stimulation parameters in this study referred to the experience of previous related studies16 and adopted individualized, adjustable parameter ranges: 1 min of electrical stimulation followed by 1 min of rest, cycled 15 times; pulse width 300 us; frequency 20 Hz; starting current intensity 0.5 mA, which was remotely adjusted by the physician during treatment according to patient tolerance until the patient's maximum tolerable level was reached (the strongest comfortable sensation without pain) rather than using fixed parameters. Sham taVNS used the same waveform with a fixed current intensity of 0.1 mA, and all other parameters were identical to taVNS.

3.7. Intervention and Control

After recruitment and screening, eligible patients were randomly assigned in a 1:1 ratio to two groups. The investigators set the parameters according to the treatment design described above and delivered continuous taVNS or sham taVNS single-blind treatment intervention for 12 weeks via remotely transmitted virtual electrical signals through the APP, once daily; the 12-week treatment was followed by a 24-week total study period. Using 4-week cycles, group A patients received taVNS, sham taVNS, and taVNS in alternating sequence, one cycle each, specifically weeks 1-4 taVNS, weeks 5-8 sham taVNS, and weeks 9-12 taVNS. Group B patients first received sham taVNS for one cycle, then two consecutive cycles of taVNS, specifically weeks 1-4 sham taVNS and taVNS from week 5 through the end of the 12-week treatment period. Blood and stool samples were collected from all participants at the end of weeks 0, 4, 8, 12, and 24, and data were gathered for follow-up assessment Figure 2.

3.8. Outcome Measures

3.8.1. Primary outcome measure: The change in CDAI score from baseline to weeks 4, 8, 12, and 24, and the number and proportion of subjects achieving clinical response (treatment efficacy) and clinical remission at weeks 4, 8, and [12].

Clinical remission was defined as a CDAI score <150, and clinical response or treatment efficacy was defined as a post-treatment CDAI score decrease of >=100 from baseline.

3.8.2. Secondary outcome measures: Including CRP, ESR, and FC levels; heart rate variability (HRV) analysis; the Inflammatory Bowel Disease Questionnaire (IBDQ); and the Hospital Anxiety and Depression Scale (HADS).

Heart rate variability (HRV): A 24-hour HRV recording was acquired using a dynamic ECG monitor (Shijiazhuang Hanwei Medical Equipment Co., Ltd., China; No. 1807030211) and analyzed with NHolter_Main 1.0.5.1 software (Hanlin Jingwei Technology Co., Ltd., China). The LF/HF ratio was used to describe the sympathovagal balance, where a decrease indicates alleviation of excessive sympathetic activation and improvement of autonomic balance. The normalized high-frequency power HFnu = HF/(LF+HF) was used to describe the proportion of parasympathetic (vagal) tone within the sympathovagal balance, where an increase indicates a relative enhancement of parasympathetic tone. In this study, LF/HF and HFnu were interpreted jointly: only when the decrease in LF/HF was accompanied by an increase in HFnu in the same direction was parasympathetic tone deemed relatively enhanced.

The IBDQ21 comprises 32 items covering four dimensions: bowel symptoms (10 items), systemic symptoms (5 items), emotional function (12 items), and social function (5 items). The total score ranges from 32 to 224, with higher scores indicating better quality of life.

The HADS22 comprises 14 items, of which 7 assess anxiety and 7 assess depression. For each subscale, a score of 0-7 is normal, 8-10 is borderline, and 11-21 is abnormal.

Subjects completed the IBDQ and HADS questionnaires at baseline and at weeks 4, 8, 12, and 24.

3.9. Safety Assessment

The overall safety of the treatment for subjects was assessed. Safety-related indicators included adverse events (AEs) and adverse reactions (ARs), as well as their incidence and time of occurrence.

4. Statistical Analysis

Statistical analyses were performed using SPSS 27.0 (IBM, Chicago, IL, USA). The Shapiro-Wilk test was used to assess the normality of each variable. The results showed that CDAI, CRP, ESR, FC, IBDQ, HADS, and LF/HF at some time points deviated from a normal distribution (Shapiro-Wilk $P<0.05$), and given the small sample size (18-20 per group) and dropouts, efficacy

indicators were uniformly described by the median (interquartile range) M (Q1, Q3); only continuous variables meeting normality, such as height, weight, and BMI, were described by mean ± standard deviation, and categorical variables by frequency (percentage). Because this trial used a crossover design, the core analysis was a within-subject paired comparison (pairing each subject's change during the true-stimulation period with that during the sham-stimulation period), using the Wilcoxon signed-rank test, with effect sizes reported as the Hodges-Lehmann location shift and its 95% CI and the rank-biserial correlation (rrb); Holm's step-down procedure was used to adjust for the 8 primary efficacy indicators (CDAI, CRP, ESR, FC, IBDQ, HADS, LF/HF, HFnu). Between-group comparisons of baseline and absolute values at each time point uniformly used the Mann-Whitney U test; given only 20 subjects per group and limited power of the normality test, the two-independent-sample t-test was used as a sensitivity analysis; binary variables used Fisher's exact test, and multi-category variables (Montreal classification) used the Fisher-Freeman-Halton exact test for Rx2 contingency tables. Between-group comparisons of changes (Delta = time point - baseline) used the Mann-Whitney U test, with baseline-adjusted rank ANCOVA (Quade's method) using baseline as a covariate to assess the effect of baseline imbalance on between-group Delta; carryover effects were tested using the Hills-Armitage method (group comparison of the sum of W4 and W8). Within-group comparisons of each follow-up time point (weeks 4, 8, 12, 24) versus baseline used the Wilcoxon signed-rank test, and the Friedman test was used to evaluate the overall time effect across the five time points. All efficacy indicators were analyzed concurrently with ITT and PP; because LOCF in the within-subject crossover paired analysis would make the missing-period change of dropouts constantly zero and systematically bias the within-group paired comparison, the per-protocol (PP) complete-case analysis was used, while the remaining descriptive between-time-point comparisons used ITT (LOCF) as the primary inference basis and PP as supportive verification. Between-group comparisons of clinical remission and response rates used Fisher's exact test. A P value <0.05 was considered statistically significant.

5. Results

5.1. General Patient Characteristics

A total of 40 patients were randomly assigned, with 20 in group A and 20 in group B. Among them, 37 (92.5%) completed the assessment and were included in the PP analysis set (19 in group A and 18 in group B); one subject in group A withdrew at week 6 because of poor disease control requiring a treatment regimen adjustment, and two subjects in group B withdrew at week 7 to switch to biologics because of inadequate efficacy; all 3 dropouts occurred before week 8, and the ITT analysis set used LOCF to impute their missing values. The two groups showed no statistically significant differences in age, sex, height, weight, BMI, disease duration, history of abdominal surgery, prior medication, Montreal classification (age at diagnosis, disease location, disease behavior, perianal disease), disease activity grade, or baseline CDAI, CRP, ESR, FC, IBDQ, and HADS (all P>0.05). (Table 1).

Table 1: Comparison of baseline characteristics between the two groups. Continuous variables are presented as mean±SD or M (Q1, Q3) according to the Shapiro-Wilk normality test; given only 20 subjects per group and limited power of the normality test, between-group comparisons uniformly used the Mann-Whitney U test, with the two-independent-sample t-test as a sensitivity analysis (yielding consistent conclusions). Binary variables used Fisher's exact test, and multi-category variables (Montreal classification: age at diagnosis, disease location, disease behavior) used the Fisher-Freeman-Halton exact test for Rx2 contingency tables.

Characteristics	Patients (n=40)		P value
	group A (n=20)	group B (n=20)	
Age (year), M(Q1,Q3)	32.5 (25.2, 48.5)	35.0 (24.5, 50.2)	0.718
Sex, n (%)			
Male, n (%)	11 (55%)	13 (65%)	0.748
Female, n (%)	9 (45%)	7 (35%)	
BMI (kg/m²), Mean±SD	19.5 ± 2.4	20.0 ± 2.2	0.502
Duration of CD (year), M(Q1,Q3)	5.0 (4.0, 7.2)	5.0 (3.0, 8.0)	0.913
Abdominal operation History, n (%)	7 (35%)	6 (30%)	1.000
Montreal classification			
Age classification			0.831
A1 (≤16), n (%)	1 (5%)	2 (10%)	
A2 (17–40), n (%)	14 (70%)	13 (65%)	
A3 (>40), n (%)	5 (25%)	5 (25%)	
Location			0.596
L1 (ileal), n (%)	8 (40%)	5 (25%)	
L2 (colonic), n (%)	3 (15%)	4 (20%)	
L3 (ileocolonic), n (%)	9 (45%)	11 (55%)	
Behavior			0.433
B1(inflammatory), n (%)	10 (50%)	6 (30%)	
B2 (stricturing), n (%)	8 (40%)	11 (55%)	
B3 (penetrating), n (%)	2 (10%)	3 (15%)	
Perianal disease, n (%)	14 (70%)	15 (75%)	1.000

Previous treatment			
5-ASA, n (%)	3 (15%)	5 (25%)	0.695
Biologics, n (%)	17(85%)	15(75%)	0.695
CDAI, M(Q1,Q3)	207.4 (192.4, 231.6)	195.7 (181.8, 209.4)	0.148
CRP (mg/L), M(Q1,Q3)	14.4 (7.9, 22.0)	13.8 (7.0, 19.6)	0.565
ESR (mm/h), M(Q1,Q3)	44.5 (26.8, 55.8)	32.0 (22.2, 44.5)	0.096
FC (μg/g), M(Q1,Q3)	89.5 (67.0, 104.0)	62.5 (53.8, 88.2)	0.072
IBDQ score, M(Q1,Q3)	168.0 (156.0, 177.2)	175.5 (165.5, 187.0)	0.102
HADS score, M(Q1,Q3)	7.0 (5.0, 9.2)	6.5 (5.0, 8.0)	0.583
LF/HF, M(Q1,Q3)	2.46 (2.10, 2.75)	2.30 (2.04, 2.57)	0.636
HFnu, M(Q1,Q3)	0.29 (0.27, 0.32)	0.30 (0.28, 0.33)	0.636

5.2. Within-Subject Crossover Paired Comparison of True versus Sham Stimulation

Given the crossover design of this trial, each subject received both a true-stimulation period and a sham-stimulation period. The change during the true-stimulation period (Deltatrue) and the change during the sham-stimulation period (Deltasham) were calculated for each subject and paired for the Wilcoxon signed-rank test, with effect sizes reported as the Hodges-Lehmann location shift and its 95% confidence interval (CI) and the rank-biserial correlation (rrb). The results showed that, relative to the sham period, the true-stimulation period produced significant median differences of -29.4 points for CDAI (Hodges-Lehmann -25.1, 95% CI -35.9 to -8.6, rrb=-0.54, P=0.004), -3.5 mg/L for CRP (-5.3, 95% CI -9.5 to -1.5, rrb=-0.50, P=0.008), -17.0 mm/h for ESR (-17.0, 95% CI -28.5 to -5.0, rrb=-0.49, P=0.009), -0.60 for LF/HF (Hodges-Lehmann -0.70, 95% CI -1.08 to -0.26, rrb=-0.56, P=0.002), and +0.06 for HFnu (+0.07, 95% CI +0.02 to +0.11, rrb=+0.57, P=0.002); these differences remained statistically significant after Holm correction (adjusted P = 0.021, 0.038, 0.038, 0.016, and 0.016, respectively). The crossover paired differences for FC (P=0.114), HADS (adjusted P=0.082), and IBDQ (P=0.920) were not statistically significant. These findings indicate that taVNS, relative to sham, reduces disease activity and inflammation levels and shifts the sympathovagal balance toward vagal predominance. Carryover-effect testing (Hills-Armitage method, between-group Mann-Whitney U of the sum of W4 and W8) yielded P>0.2 for all indicators, with no significant carryover effect detected. Between-group comparisons of changes (Delta = time point - baseline) showed that at week 4 (group A true vs. group B sham), the reductions in CDAI, CRP, and ESR were significantly greater in group A than in group B (CDAI: -19.8 vs. +9.3, P=0.018; CRP: -2.3 vs. +0.4 mg/L, P=0.014; ESR: -17.0 vs. -1.5 mm/h, P=0.007), and the increase in HF was also significantly greater (group A vs. group B: +398.5 vs. -55.5 ms2, P<0.001); after Quade baseline correction, the between-group Delta differences for CRP (P=0.016), ESR (P=0.029), and HF (P<0.001) remained significant, whereas CDAI became borderline (P=0.054).

5.3. CDAI Score

The ITT analysis showed no statistically significant between-group difference in baseline CDAI score. In group A, the CDAI score decreased from baseline at week 12 (192.3 [172.8, 209.8], P=0.014) and week 24 (189.9 [166.6, 209.8], P=0.011), remaining significant after Holm correction (adjusted P=0.043 for both); the changes at week 4 (P=0.070, adjusted P=0.139) and week 8 (sham period, P=0.956) were not significant. For group B, the P values for comparison with baseline at each time point were 0.076 (week 4), 0.522 (week 8), 0.388 (week 12), and 0.048 (week 24); only week 24 [173.6 (167.8, 205.7)] reached unadjusted statistical significance, which did not persist after correction (adjusted P=0.194). No statistically significant between-group differences were observed at any time point (week 4 P=0.231, week 8 P=0.192, week 12 P=0.779, week 24 P=0.799). The Friedman test indicated a decrease in CDAI over time in both groups (group A P<0.001, group B P=0.002) (Figure 3A).

The PP analysis showed no statistically significant between-group difference in baseline CDAI score [group A 205.9 (191.8, 231.9), group B 194.9 (181.2, 205.8), P=0.111]. In group A, CDAI decreased from baseline at week 12 (P=0.018) and week 24 (P=0.014) but did not reach significance after Holm correction (adjusted P=0.056 for both); the changes at week 4 (P=0.087) and week 8 (sham period, P=0.953) were not significant. In group B, only week 24 showed a significant decrease from baseline [171.1 (164.7, 191.3), P=0.001], which remained significant after correction (adjusted P=0.004); the other time points (week 4 P=0.212, week 8 P=0.119, week 12 P=0.074) were not significant. No statistically significant between-group differences were observed at any time point (week 4 P=0.374, week 8 P=0.066, week 12 P=0.480, week 24 P=0.518). The Friedman test indicated a decrease in CDAI over time in both groups (group A P=0.001, group B P<0.001) (Figure 3B).

The clinical remission and response rates in both groups were generally low. The ITT analysis showed that the clinical remission rate in group A was 0% at weeks 4, 8, and 12 and 10% (2/20) at week 24 (follow-up period), and the clinical response rate was 0% at all time points. In group B, the clinical remission rate was 0% at all time points, and the clinical response rate was 5% (1/20) at week 12 (taVNS period) and 5% (1/20) at week 24 (follow-up period). The PP analysis (group A n=19, group B n=18) showed consistent direction: only group A reached clinical remission of 10.5% (2/19) at week 24, and the rest were 0%; only group B reached clinical response of 5.6% (1/18) at weeks 12 and 24, and the rest were 0%. No statistically significant between-group differences in clinical remission or response rates were observed at any time point (Fisher's exact test, all P>0.05). During the taVNS intervention period (group A weeks 4 and 12, group B weeks 8 and 12), no subject met the criteria for clinical remission or response (Figures 3C-F).

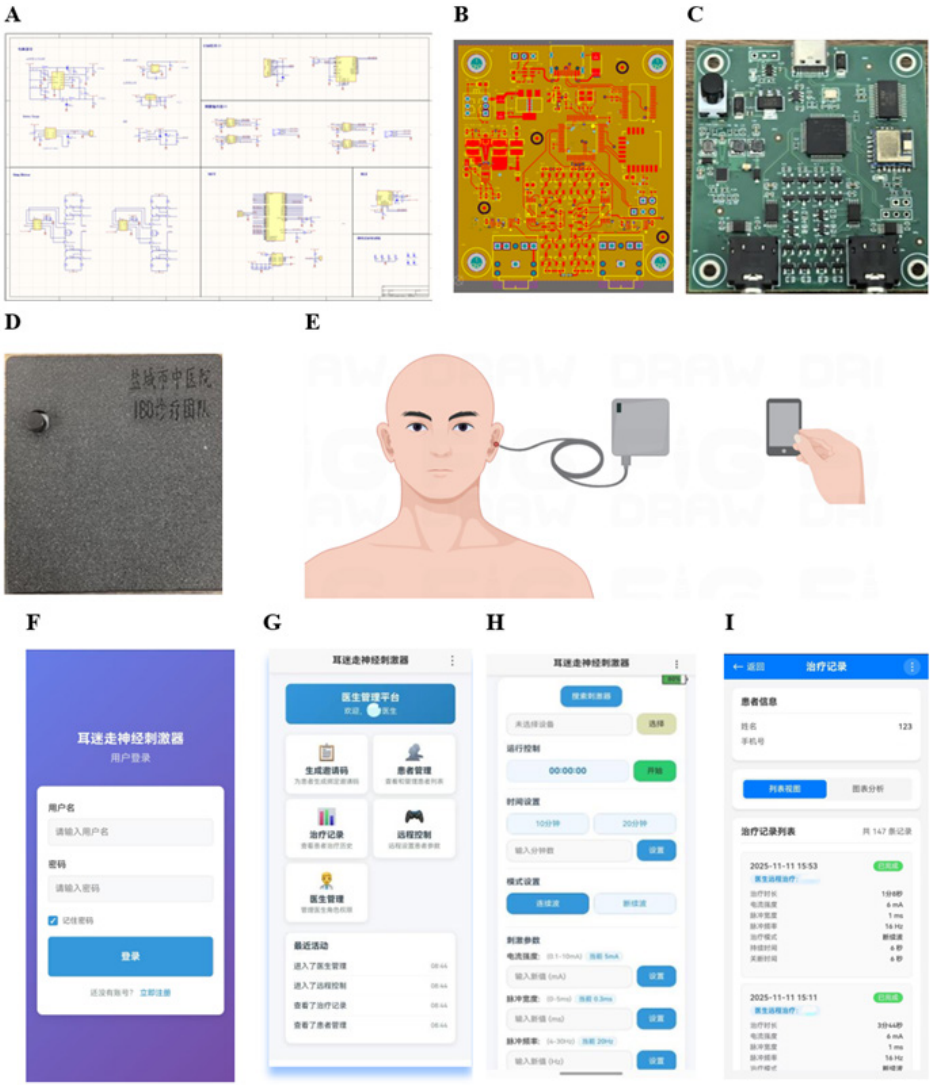


Figure 1: Schematic of the remotely controlled auricular vagus nerve stimulator and its companion mobile APP. A, circuit diagram; B, circuit layout; C, physical circuit photograph; D, device appearance; E, connection diagram for use; F, APP login interface; G, physician main interface; H, treatment parameter setting interface; I, physician-end patient treatment record table (example).

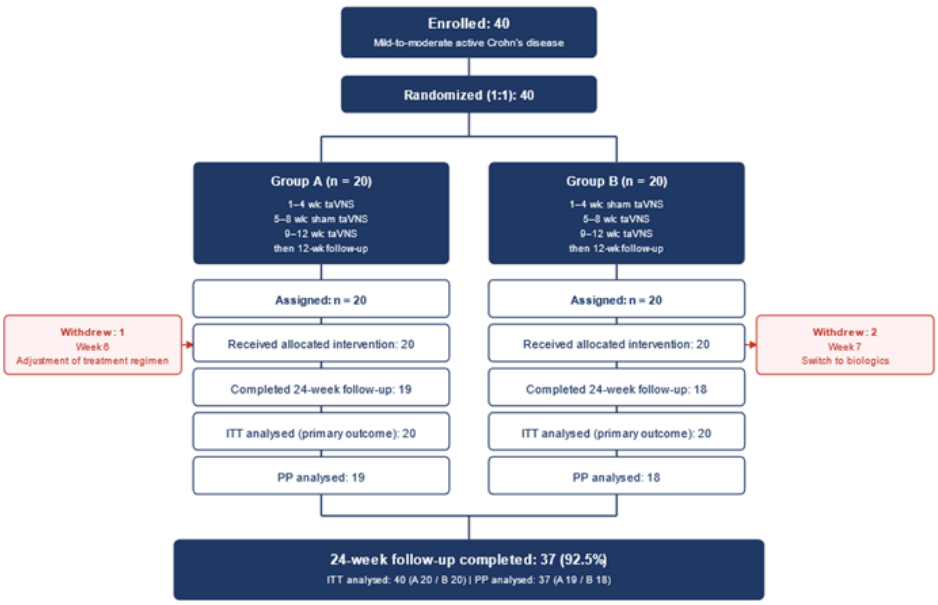


Figure 2: Schematic of the experimental grouping and flow. A total of 40 patients with mild-to-moderate active Crohn's disease were randomly assigned in a 1:1 ratio to group A (n=20) and group B (n=20). The intervention sequence for group A was taVNS-sham taVNS-taVNS (weeks 1-4, weeks 5-8 sham, weeks 9-12 taVNS), and for group B it was sham taVNS-taVNS-taVNS (weeks 1-4 sham, weeks 5-12 taVNS), with 4-week cycles and a 12-week treatment followed by a 12-week follow-up. Thirty-seven patients (92.5%) completed all five time-point assessments and entered the per-protocol (PP) analysis set (19 in group A, 18 in group B); 3 patients withdrew before week 8, and the intention-to-treat (ITT) analysis used the last-observation-carried-forward (LOCF) method for imputation. Primary and secondary efficacy endpoints were collected at weeks 0, 4, 8, 12, and 24.

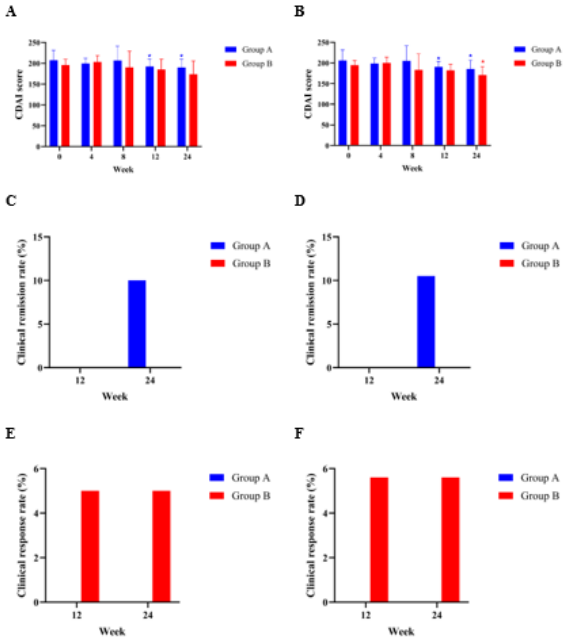


Figure 3: Changes in CDAI scores over time and post-treatment clinical remission and response rates in the two groups. A, CDAI score change, intention-to-treat (ITT) analysis (group A n=20, group B n=20); B, CDAI score change, per-protocol (PP) analysis (group A n=19, group B n=18); C, clinical remission rate, ITT analysis (group A n=20, group B n=20); D, clinical remission rate, PP analysis (group A n=19, group B n=18); E, clinical response rate, ITT analysis (group A n=20, group B n=20); F, clinical response rate, PP analysis (group A n=19, group B n=18). Data in panels A and B are presented as M (Q1, Q3); all between-group comparisons at each time point had P>0.05. The within-group direction of CDAI change was consistent across ITT and PP analyses; the improvement in group B at week 24 reached significance only after multiple-comparison correction in the PP analysis. Clinical remission (CDAI<150); clinical response (CDAI decrease >=100 from baseline). Between-group comparisons at each time point used Fisher's exact test, all P>0.05. *P<0.05, **P<0.01, ***P<0.001.

5.4. CRP

No statistically significant between-group difference in baseline CRP was observed. The ITT analysis showed that in group A, CRP decreased from baseline at week 4 (taVNS period) (8.4 [4.6, 15.9] mg/L, P=0.001), week 12 (taVNS period) (8.0 [5.0, 18.0] mg/L, P=0.008), and week 24 (follow-up period) (7.2 [3.2, 16.1] mg/L, P=0.008), remaining significant after Holm correction (adjusted P=0.004, 0.025, 0.025); the change at week 8 (sham period) (12.5 [7.5, 19.1] mg/L, P=0.368) was not significant. In group B, CRP decreased from baseline at week 24 (follow-up period) (5.5 [2.8, 11.9] mg/L, P=0.027) but did not remain significant after correction (adjusted P=0.107); weeks 4 (sham period, P=0.452), 8 (P=0.368), and 12 (P=0.097) were not significant. No statistically significant between-group differences were observed at any time point (week 4 P=0.383, week 8 P=0.314, week 12 P=0.529, week 24 P=0.583). The Friedman test showed an overall change in CRP over time (ITT: group A P=0.002, group B P=0.005) (Figure 4A).

The PP analysis showed that in group A, CRP decreased from baseline at week 4 (8.1 [4.5, 13.6] mg/L, P<0.001), week 12 (7.6 [4.8, 15.9] mg/L, P=0.008), and week 24 (4.3 [3.2, 14.8] mg/L, P=0.007), remaining significant after correction (adjusted P=0.004, 0.021, 0.021); the change at week 8 (sham period) (11.7 [6.8, 18.6] mg/L, P=0.395) was not significant. In group B, CRP decreased from baseline at week 12 (5.0 [3.3, 8.4] mg/L, P=0.004) and week 24 (4.8 [2.4, 8.2] mg/L, P<0.001), remaining significant after correction (adjusted P=0.012, 0.001); weeks 4 (sham period, P=0.899) and 8 (P=0.067) were not significant. No statistically significant between-group differences were observed at any time point (week 4 P=0.558, week 8 P=0.150, week 12 P=0.298, week 24 P=0.343). The Friedman test showed an overall change in CRP over time (PP: group A P<0.001, group B P<0.001) (Figure 4B).

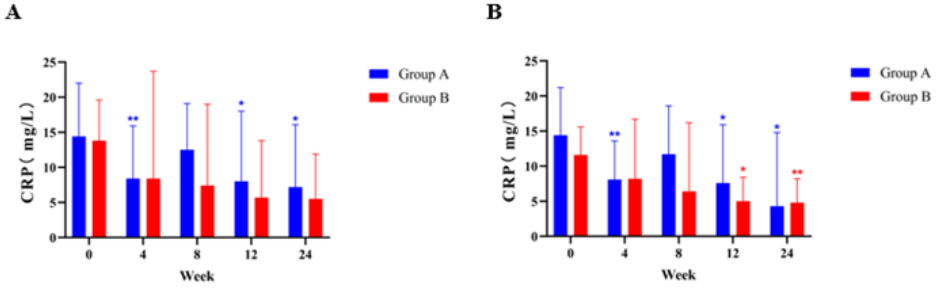


Figure 4: Changes in serum C-reactive protein (CRP) levels over time in the two groups. A, intention-to-treat (ITT) analysis; B, per-protocol (PP) analysis. All between-group comparisons at each time point had P>0.05. The within-group decreases in group A at weeks 4, 12, and 24 (ITT and PP analyses) and in group B at weeks 12 and 24 (PP analysis) remained significant after multiple-comparison correction. *P<0.05, **P<0.01, ***P<0.001.

5.5. ESR

No statistically significant between-group difference in baseline ESR was observed. The ITT analysis showed that in group A, ESR decreased from baseline at week 4 (24.5 [15.5, 40.5] mm/h, P=0.002), week 8 (27.0 [15.8, 62.2] mm/h, P=0.036), week 12 (18.0 [12.0, 34.0] mm/h, P<0.001), and week 24 (15.0 [7.0, 33.2] mm/h, P<0.001), all remaining significant after Holm correction (adjusted P=0.003, 0.036, <0.001, <0.001). In group B, ESR decreased from baseline at week 8 (18.0 [13.8, 37.2] mm/h, P=0.006), week 12 (20.5 [11.8, 24.2] mm/h, P=0.002), and week 24 (17.5 [7.0, 29.2] mm/h, P=0.001), remaining significant after correction (adjusted P=0.011, 0.007, 0.004); the change at week 4 (sham period) (29.0 [23.0, 46.2] mm/h, P=0.784) was not significant. No statistically significant between-group differences were observed at any time point (week 4 P=0.383, week 8 P=0.102, week 12 P=0.758, week 24 P=0.862). The Friedman test showed an overall change in ESR over time (ITT: both group A and group B P<0.001) (Figure 5A).

The PP analysis showed that in group A, ESR decreased from baseline at week 4 (25.0 [15.0, 41.0] mm/h, P=0.002), week 8 (27.0 [15.5, 62.5] mm/h, P=0.043), week 12 (15.0 [12.0, 35.0] mm/h, P<0.001), and week 24 (14.0 [7.0, 36.5] mm/h, P<0.001), all remaining significant after correction (adjusted P=0.005, 0.043, <0.001, <0.001). In group B, ESR decreased from baseline at week 8 (15.5 [13.2, 38.2] mm/h, P=0.001), week 12 (17.0 [11.2, 24.0] mm/h, P<0.001), and week 24 (14.5 [7.0, 27.0] mm/h, P<0.001), remaining significant after correction (adjusted P=0.002, 0.001, 0.001); the change at week 4 (sham period) (29.0 [23.0, 48.8] mm/h, P=0.523) was not significant. No statistically significant between-group differences were observed at any time point (week 4 P=0.443, week 8 P=0.075, week 12 P=0.620, week 24 P=0.753). The Friedman test showed an overall change in ESR over time (PP: both group A and group B P<0.001) (Figure 5B).

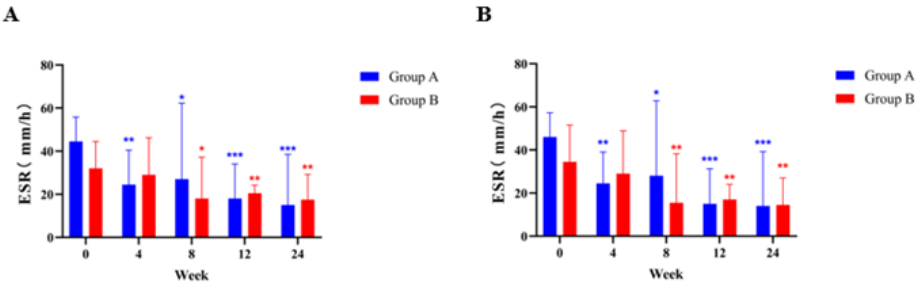


Figure 5: Changes in erythrocyte sedimentation rate (ESR) over time in the two groups. A, intention-to-treat (ITT) analysis; B, per-protocol (PP) analysis. Data are presented as M (Q1, Q3). All between-group comparisons at each time point had P>0.05. The ITT and PP analyses yielded consistent directional conclusions. *P<0.05, **P<0.01, ***P<0.001.

5.6. Fecal Calprotectin

No statistically significant between-group difference in baseline FC was observed. The ITT analysis showed that in group A, FC decreased from baseline at week 12 (52.0 [46.8, 79.2] ug/g, P=0.007) and week 24 (39.5 [29.0, 54.8] ug/g, P<0.001), remaining significant after Holm correction (adjusted P=0.022, 0.004); weeks 4 (P=0.096) and 8 (sham period, P=0.956) were not significant. In group B, FC decreased from baseline at week 24 (52.0 [42.0, 69.2] ug/g, P=0.024) but did not remain significant after correction (adjusted P=0.096); weeks 4 (sham period, P=0.956), 8 (P=0.475), and 12 (P=0.097) were not significant. No statistically significant between-group differences were observed at any time point (week 4 P=1.000, week 8 P=0.201, week 12 P=0.820, week 24 P=0.081). The Friedman test showed an overall change in FC over time (ITT: group A P<0.001, group B P=0.023) (Figure 6A).

The PP analysis showed that in group A, FC decreased from baseline at week 12 (53.9 [47.5, 80.5] ug/g, P=0.014) and week 24 (40.0 [30.0, 59.5] ug/g, P=0.002), remaining significant after correction (adjusted P=0.042, 0.008); weeks 4 (P=0.167) and 8 (sham period, P=0.651) were not significant. In group B, FC decreased from baseline at week 24 (50.5 [42.0, 59.5] ug/g, P=0.001), remaining significant after correction (adjusted P=0.005); week 12 (55.9 [46.8, 63.0] ug/g, P=0.016) did not remain significant after correction (adjusted P=0.048), and weeks 4 (sham period, P=0.495) and 8 (P=0.196) were not significant. No statistically significant between-group differences were observed at any time point (week 4 P=0.730, week 8 P=0.075, week 12 P=0.822, week 24 P=0.189). The Friedman test showed an overall change in FC over time (PP: group A P<0.001, group B P=0.007) (Figure 6B).

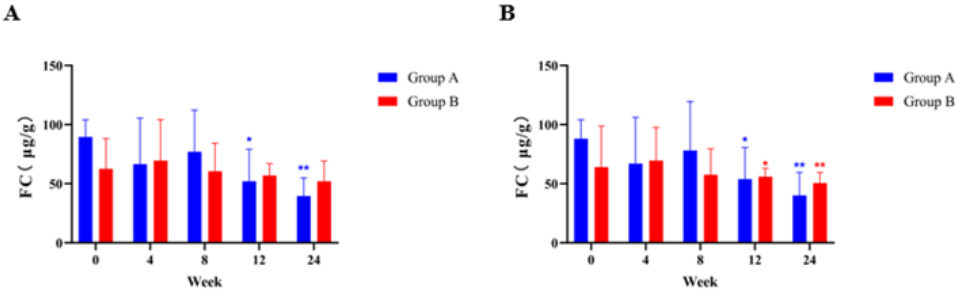


Figure 6: Changes in fecal calprotectin (FC) levels over time in the two groups. A, intention-to-treat (ITT) analysis; B, per-protocol (PP) analysis. Data are presented as M (Q1, Q3). All between-group comparisons at each time point had P>0.05. The within-group decreases in group A at weeks 12 and 24 and in group B at week 24 (PP analysis) remained significant after multiple-comparison correction. *P<0.05, **P<0.01, ***P<0.001.

5.7. IBDQ

No statistically significant between-group difference in baseline IBDQ score was observed; however, in the PP analysis set the between-group baseline difference was significant [group A 166.0 (156.0, 176.5), group B 176.5 (168.0, 187.0), P=0.017]. The ITT analysis showed that in group A, the week-12 score (172.5 [164.0, 187.5], P=0.040) differed significantly from baseline but did not reach significance after Holm correction (adjusted P=0.161); weeks 4 (P=0.418), 8 (sham period, P=0.090), and 24 (P=0.130) were not significant. In group B, no time point differed significantly from baseline (week 4 P=0.679, week 8 P=0.869, week 12 P=0.546, week 24 P=0.498). No statistically significant between-group differences were observed at any time point (week 4 P=0.583, week 8 P=0.968, week 12 P=0.461, week 24 P=0.602). The Friedman test showed an overall change in IBDQ score over time (ITT: group A P=0.027, group B P=0.084) (Figure 7A).

The PP analysis showed that in group A, the week-12 score (172.0 [164.0, 188.0], P=0.008) differed significantly from baseline and remained significant after Holm correction (adjusted P=0.031); weeks 8 (sham period, P=0.032) and 24 (P=0.051) did not reach significance after correction (adjusted P=0.097, 0.101), and week 4 (P=0.212) was not significant. In group B, no time point differed significantly from baseline (week 4 P=0.818, week 8 P=0.304, week 12 P=0.142, week 24 P=0.119). No statistically significant between-group differences were observed at any time point (week 4 P=0.233, week 8 P=0.518, week 12 P=0.178, week 24 P=0.271). The Friedman test showed an overall change in IBDQ score over time (PP: group A P=0.011, group B P=0.046) (Figure 7B).

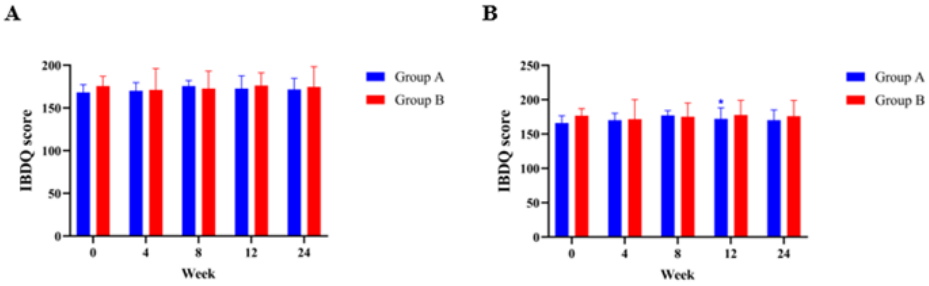


Figure 7: Changes in Inflammatory Bowel Disease Questionnaire (IBDQ) scores over time in the two groups. A, intention-to-treat (ITT) analysis; B, per-protocol (PP) analysis. Data are presented as M (Q1, Q3), with higher scores indicating better quality of life. All between-group comparisons at each time point had P>0.05. Only the within-group improvement in group A at week 12 (PP analysis) remained significant after multiple-comparison correction. *P<0.05, **P<0.01, ***P<0.001.

5.8. HRV

HRV must be interpreted in combination with LF/HF and HFnu. LF/HF reflects the sympathovagal balance, and HFnu = HF/(LF+HF) reflects the proportion of parasympathetic (vagal) tone; a decrease in LF/HF accompanied by an increase in HFnu is interpreted as a relative enhancement of parasympathetic tone. No statistically significant between-group differences in baseline LF/HF or HFnu were observed (ITT P=0.640; PP P=0.685); however, baseline imbalance existed in the absolute frequency-domain values of HRV (ITT: LF P=0.014, HF P=0.013; PP: LF P=0.009, HF P=0.006). Same-time-point between-group comparisons showed that at week 4 (group A true vs. group B sham) and week 8 (group A sham vs. group B true), the between-group differences in LF/HF and HFnu were both significant (ITT P=0.004; PP week 4 P=0.009, week 8 P=0.001), and the directions were opposite, indicating that the autonomic balance changes during the true- and sham-stimulation periods were in opposite directions. At week 12, both groups received true stimulation, and the between-group difference was not significant in the ITT analysis (P=0.096) but reached significance in the PP analysis (P=0.026); at week 24 (follow-up period), no significant between-group difference was observed (ITT P=0.231, PP P=0.075). Within-group comparisons versus baseline after Holm correction showed that in group A at week 4 (taVNS period), LF/HF decreased from 2.46 (2.10, 2.75) to 1.99 (1.70, 2.24) and HFnu increased from 0.29 (0.27, 0.32) to 0.33 (0.31, 0.37), but neither change was significant (P=0.312, 0.261); at week 8 (transition to sham period), both returned synchronously (P=0.430, 0.349); only at week 12 did the LF/HF decrease remain significant (ITT adjusted P=0.026, PP adjusted P=0.049), whereas the HFnu change did not remain significant after correction (ITT P=0.030, adjusted P=0.118). In group B, at week 4 (sham period) LF/HF increased to 2.49 (2.23, 2.64) and HFnu decreased to 0.29 (0.27, 0.31) (P=0.154, 0.177); after transitioning to taVNS at week 8, LF/HF significantly decreased to 1.87 (1.54, 2.26) and HFnu significantly increased to 0.35 (0.31, 0.39) (ITT P=0.021, 0.015; adjusted P=0.043, 0.031), and these changes persisted at weeks 12 and 24 and remained significant after correction (LF/HF: P<0.001, 0.006; adjusted P=0.002, 0.019; HFnu: P<0.001, 0.006; adjusted P<0.001, 0.017). The Friedman test showed that LF/HF and HFnu changed significantly over time in both groups (all P<0.001). Group A's HF was significantly increased from baseline at W4, W12, and W24 (all P<0.001), and group B's HF was significantly increased at W8, W12, and W24 (P=0.004, <0.001, 0.015) (Figure 8).

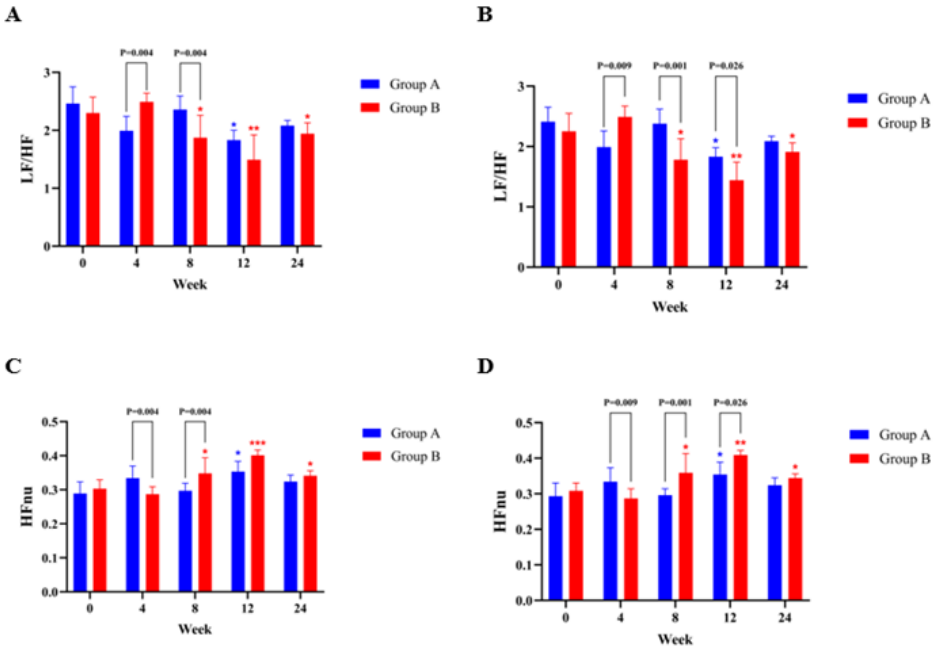


Figure 8: Changes in heart rate variability indices in the two groups. A, ITT analysis of LF/HF ratio; B, PP analysis of LF/HF ratio; C, ITT analysis of HFnu; D, PP analysis of HFnu. A decrease in LF/HF accompanied by an increase in HFnu indicates a shift of the sympathovagal balance toward vagal predominance. The ITT analysis showed significant between-group differences in LF/HF and HFnu at weeks 4 and 8 (both P=0.004); the PP analysis showed significant between-group differences in LF/HF and HFnu at weeks 4, 8, and 12 (P=0.009, 0.001, 0.026); the between-group difference at week 24 was not significant (ITT P=0.231, PP P=0.075). *P<0.05, **P<0.01, ***P<0.001.

5.9. HADS

No statistically significant between-group difference in baseline HADS score was observed. The ITT analysis showed that in group A, the week-12 score (6.0 [4.0, 7.0], P=0.014) differed significantly from baseline but did not reach significance after Holm correction (adjusted P=0.056); weeks 4 (P=0.212), 8 (sham period, P=0.720), and 24 (P=0.073) were not significant. In group B, no time point differed significantly from baseline (week 4 P=0.303, week 8 P=0.528, week 12 P=0.561, week 24 P=0.442). No statistically significant between-group differences were observed at any time point (week 4 P=0.512, week 8 P=0.512, week 12 P=1.000, week 24 P=0.820); the Friedman test showed an overall change in HADS score over time (ITT: group A P=0.036, group B P=0.253) (Figure 9A).

The PP analysis showed that in group A, the week-12 score (6.0 [4.0, 7.0], P=0.027) differed significantly from baseline but did not reach significance after correction (adjusted P=0.107); the remaining time points (week 4 P=0.329, week 8 P=0.426, week 24 P=0.130) were not significant. In group B, no time point differed significantly from baseline (week 4 P=0.787, week 8 P=0.091, week 12 P=0.110, week 24 P=0.093). No statistically significant between-group differences were observed at any time point (week 4 P=0.988, week 8 P=0.159, week 12 P=0.461, week 24 P=0.327). The Friedman test showed no significant overall change in HADS score over time (PP: group A P=0.053, group B P=0.125) (Figure 9B).

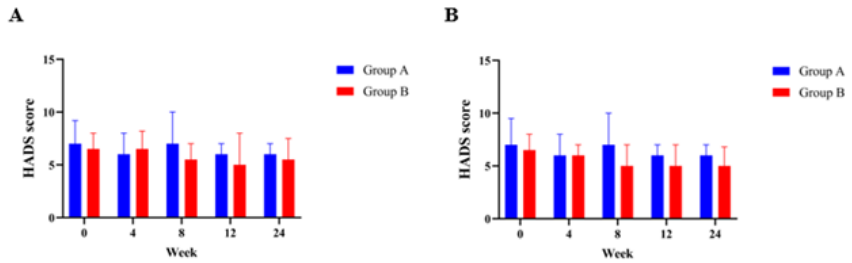


Figure 9: Changes in Hospital Anxiety and Depression Scale (HADS) scores over time in the two groups. A, intention-to-treat (ITT) analysis; B, per-protocol (PP) analysis. Data are presented as M (Q1, Q3), with higher scores indicating more severe anxiety and depression. All between-group comparisons at each time point had P>0.05. No within-group result at any time point reached significance after multiple-comparison correction. *P<0.05, **P<0.01, ***P<0.001.

5.10. Safety and Tolerability Assessment

A total of 40 subjects were randomly assigned in this study, and 37 (92.5%) completed the 24-week follow-up; one subject in group A withdrew at week 6 because of worsening disease activity requiring a change in medication regimen, and two subjects in group B withdrew at week 7 to switch to biologics because of inadequate efficacy. The most common adverse event was mild auricular clamp pain during treatment (one case in group A, two in group B), which resolved after local pressing and kneading. No serious adverse events occurred during the study. Overall, taVNS showed good safety and tolerability within the stimulation intensity range from a starting 0.5 mA up to each patient's maximum painless tolerable level.

6. Discussion

This study employed a randomized controlled design to systematically evaluate the efficacy and safety of taVNS-based on an independently developed remotely controlled auricular vagus nerve stimulator-in patients with mild-to-moderate active CD. All efficacy outcome measures were analyzed concurrently with PP and ITT. The results showed that at the within-group level, group A exhibited significant decreases in CRP (ITT P=0.001, PP P<0.001) and ESR (ITT and PP both P=0.002) as early as week 4 of taVNS, whereas group B, which was in the sham period at the same time, showed no significant changes in these indicators (CRP: ITT P=0.452, PP P=0.899; ESR: ITT P=0.784, PP P=0.523); after group B transitioned to taVNS at week 8, ESR (ITT P=0.006, PP P=0.001) and other indicators showed sequential significant improvements, and FC decreased significantly at weeks 12 and 24 in both groups (group A PP P=0.014, 0.002; group B PP P=0.016, 0.001), suggesting a temporal correspondence between the improvement in inflammatory indicators and the taVNS stimulation assignment. The primary endpoint CDAI improved within group mainly at week 12 (ITT P=0.014, PP P=0.018) and week 24 (ITT P=0.011, PP P=0.014) in group A and at week 24 in group B (ITT P=0.048, PP P=0.001). Within-group improvement in IBDQ was seen at weeks 8 and 12 in group A (PP P=0.032, 0.008; ITT week 12 P=0.040), and HADS improved significantly only at week 12 in group A (ITT P=0.014, PP P=0.027), but neither remained significant after Holm correction. No serious adverse events were observed throughout the study.

The core theoretical basis of this study is the cholinergic anti-inflammatory pathway. After excitation of vagal efferent fibers, norepinephrine is released via the splenic nerve, prompting specific T-cell subsets in the spleen to release acetylcholine, which then binds to the alpha7nAChR on the surface of immune cells such as macrophages, inhibiting the synthesis and release of pro-inflammatory cytokines such as TNF-α, IL-1beta, and IL-69, 23. In addition, after vagal afferent fibers sense peripheral inflammatory signals, they project to the nucleus of the solitary tract and the dorsal motor nucleus of the vagus, and through central integration form a closed-loop anti-inflammatory reflex via efferent fibers24. This study provides multidimensional evidence for the activation of CAP in CD patients: CRP and ESR decreased significantly as early as week 4 of taVNS in group A, while FC decreased significantly only in the later treatment phase (weeks 12 and 24), suggesting that inflammatory control may progress gradually from systemic to local intestinal levels. At the neurophysiological level, after group B transitioned to taVNS at week 8, LF/HF decreased significantly (P=0.021) and remained significant after Holm correction (adjusted P=0.043); at week 12, when both groups received taVNS, group B showed a significant LF/HF change (P<0.001, adjusted P=0.002) and group A showed a significant LF/HF decrease (P=0.006, adjusted P=0.026); in contrast, at week 8 when group A transitioned from taVNS to sham and at week 4 when group B was in the sham period, the above ratio indicators did not show directionally consistent significant changes, providing objective neurophysiological evidence for the vagal activation effect of taVNS25. These findings connect the chain linking neural stimulation to anti-inflammatory effects and provide a mechanistic explanation for taVNS in the treatment of CD.

The data of this study are partly consistent with existing exploratory results, further suggesting the potential of neuromodulation in the treatment of CD13, 16, 26. Through the alternating true- and sham-stimulation design, after group A switched from taVNS to sham, CRP rebounded significantly at week 8 compared with week 4 (direct comparison between adjacent time points within group: P=0.009), and CDAI increased synchronously (P=0.020); whereas after group B switched from sham to taVNS, CRP (P=0.043) and ESR (P=0.011) decreased further at week 8 compared with week 4, and LF/HF decreased significantly (P<0.001). This phenomenon is consistent with the reversibility of the anti-inflammatory effect of VNS observed by Caravaca et al.27 in an animal model, and to some extent supports the reversibility and stimulation-dependence of the taVNS effect; after formal between-group change (Delta) comparison and Quade baseline correction, the between-group Delta differences for CRP, ESR, and HF at week 4 remained significant, and the carryover-effect test was not significant, so temporal effects or natural disease course fluctuations cannot fully explain the above stimulation-dependent changes. The rapid rebound of inflammatory indicators during the sham period suggests that the physiological effect of taVNS may be short-lived, supporting the necessity of daily continuous treatment. In this study, the 12-week follow-up after the end of treatment (week 24) showed that the improvement in CDAI was maintained in both groups (group A ITT P=0.011, PP P=0.014; group B ITT P=0.048, PP P=0.001), and group B's HRV LF/HF at week 24 remained significantly lower than baseline (P=0.006, adjusted P=0.019), suggesting that some of the benefits induced by taVNS have a certain persistence after treatment cessation. In contrast, the within-group improvements in IBDQ and HADS during the follow-up period were mostly not significant (IBDQ: group A ITT week 24 P=0.130, group B ITT week 24 P=0.498; HADS: group A P=0.073, group B P=0.442), indicating limited maintenance of quality-of-life and mood benefits, and longer-term or continuous-stimulation regimens need to be evaluated in the future. Furthermore, the effect of taVNS on HADS score (group A week 12 ITT P=0.014, adjusted P=0.056)

suggests that this treatment may exert both anti-inflammatory and mood-regulating effects through vagus nerve-brain axis regulation. Liu et al.28 have reported that taVNS can exert antidepressant and anti-inflammatory effects by modulating the prefrontal-limbic neural circuit, and our observations in CD patients are in the same vein, suggesting that the benefits of taVNS in IBD may extend beyond mere intestinal anti-inflammation to involve broader brain-gut-immune axis regulation.

In addition, the remotely controlled auricular vagus nerve stimulator used in this study achieved innovative breakthroughs at both the hardware and software levels. At the hardware level, the improved complementary Wilson current-mirror constant-current drive circuit and the H-bridge charge-balanced output stage ensure the precise controllability of the stimulation dose and the safety of long-term treatment from the circuit topology. At the software level, the IoT-based APP remote management system realizes a closed-loop treatment model of 'physician sets parameters -> cloud transmission -> patient performs at home -> data return -> efficacy monitoring.' This model is of great value for chronic diseases such as CD that require long-term management. In the post-COVID-19 pandemic era, telemedicine has become an important trend in chronic disease management, and the remote-control capability of taVNS can reduce unnecessary hospital visits for patients, lower medical costs, and improve treatment adherence [29].

In terms of safety, taVNS performed excellently. All subjects tolerated the set stimulation intensity range, and no one withdrew from the trial because of local discomfort or stimulation-related side effects. This is consistent with the noninvasive and low-risk nature of the technique15, 30. In this study, 0.1 mA was selected as the sham stimulation current, below the sensory threshold of the vast majority of individuals, theoretically maintaining good blinding. However, it cannot be completely ruled out that some sensitive individuals might guess their group assignment due to perception of slight skin sensation differences. Future studies could add blinding assessment sessions and consider more refined sham protocols such as lobule stimulation deviating from vagal distribution areas or blunt electrodes. Furthermore, the stimulation parameters of the remotely controlled auricular vagus nerve stimulator used in this study can be adjusted according to individual tolerance: parameters such as frequency and pulse width can be individually set for each patient by the investigator through the cloud platform, and the current intensity is progressively adjusted from 0.5 mA to the patient's maximum painless tolerable level according to tolerance, which has adapted to some extent to inter-patient differences in auricular anatomy, nerve fiber recruitment thresholds, and current penetration depth.

However, this study has the following limitations that need to be objectively considered when interpreting the results. As a single-center, small-sample (n=40) exploratory trial, the generalizability of the findings is limited. Moreover, the efficacy evaluation indicators in this study were mainly based on clinical symptoms and biomarkers, lacking direct evidence of post-treatment endoscopic mucosal healing or ultrasound assessment. As a treat-to-target goal in CD, endoscopic remission is the gold standard for assessing disease-modifying effects, and its absence is an important shortcoming of this stage of research31. The 24-week trial period can only reflect short-term efficacy; as a lifelong disease, whether taVNS can maintain long-term remission and alter the natural history of CD remains unknown. In addition, this study adopted a single-blind design, and efficacy indicators such as IBDQ and HADS are subjective scales, so the influence of expectation bias on the results cannot be completely excluded.

7. Conclusion

This study shows that taVNS based on an independently developed remotely controlled auricular vagus nerve stimulator has good safety and tolerability in patients with mild-to-moderate active CD. With the within-subject crossover paired comparison as the core evidence, taVNS significantly reduced disease activity (CDAI) and inflammatory markers (CRP, ESR) relative to sham stimulation and decreased HRV LF/HF while increasing HFnu (shifting the sympathovagal balance toward vagal predominance). As a single-center, small-sample trial, these conclusions still require validation in larger samples. The remote-control function makes this treatment model feasible for long-term home-based management, and future randomized controlled trials with larger samples, longer durations, and endoscopic endpoints and blinding assessment are warranted.

8. Funding

Yancheng Municipal Health Commission Project (NO. YK2024134), Science and Technology Planning Project of Yancheng City (NO. YCBK2025037), Postgraduate Research and Practice Innovation Program of Jiangsu Province (NO. SJCX25_1085)

9. Ethics Approval and Consent To Participate

This study was conducted in accordance with the Declaration of Helsinki (revised in 2013), and written informed consent was obtained from all patients prior to treatment. The study protocol was approved by the Ethics Committee of Yancheng TCM Hospital Affiliated to Nanjing University of Chinese Medicine (approval No. KY250413) and was registered at the Chinese Clinical Trial Registry (ChiCTR, identifier: ChiCTR2500105702). They were informed that their medical records might be utilized in future medical research. All identifiable patient information has been de-identified.

10. Consent for Publication

All authors have read and approved the final manuscript for submission and agree with its publication in the journal.

References

1. Viola A, Fiorino G, Costantino G, Fries W. Epidemiology and clinical course of late onset inflammatory bowel disease. Minerva Gastroenterol (Torino). 2024; 70(1): 52-58.
2. Gordon H, Burisch J, Ellul P. ECCO Guidelines on Extraintestinal Manifestations in Inflammatory Bowel Disease. J Crohns Colitis. 2024; 18(1): 1-37.
3. Windsor JW, Kaplan GG. Evolving Epidemiology of IBD. Curr Gastroenterol Rep. 2019; 21(8): 40.
4. Kaplan GG, Windsor JW. The four epidemiological stages in the global evolution of inflammatory bowel disease. Nat Rev Gastroenterol Hepatol. 2021; 18(1): 56-66.
5. Radi ZA, Heuvelman DM, Masferrer JL, Benson EL. Pharmacologic evaluation of sulfasalazine, FTY720, and anti-IL-12/23p40 in a TNBS-induced Crohn's disease model. Dig Dis Sci. 2011; 56(8): 2283-2291.
6. Gordon H, Minozzi S, Kopylov U. ECCO Guidelines on Therapeutics in Crohn's Disease: Medical Treatment. J Crohns Colitis. 2024; 18(10): 1531-1555.
7. Shin JH, Lee CM, Song JJ. Transcutaneous auricular vagus nerve stimulation mitigates gouty inflammation by reducing neutrophil infiltration in BALB/c mice. Sci Rep. 2024; 14(1): 25630.
8. Pikov V. Vagus Nerve Stimulation and Sacral Nerve Stimulation for Inflammatory Bowel Disease: A Systematic Review. J Transl Gastroenterol. 2023; 1(2): 94-100.
9. Borovikova LV, Ivanova S, Nardi D. Role of vagus nerve signaling in CNI-1493-mediated suppression of acute inflammation. Auton Neurosci. 2000; 85(1-3): 141-147.
10. Ji H, Rabbi MF, Labis B, Pavlov VA, Tracey KJ, Ghia JE. Central cholinergic activation of a vagus nerve-to-spleen circuit alleviates experimental colitis. Mucosal Immunol. 2014; 7(2): 335-347.
11. Zi S, Li J, Liu L, Liu F. Cholinergic anti-inflammatory pathway and its role in treatment of sepsis. Zhong Nan Da Xue Xue Bao Yi Xue Ban. 2020; 45(1): 68-73.
12. D'Haens G, Eberhardson M, Cabrijan Z. Neuroimmune Modulation Through Vagus Nerve Stimulation Reduces Inflammatory Activity in Crohn's Disease Patients: A Prospective Open-label Study. J Crohns Colitis. 2023; 17(12): 1897-1909.
13. Sinniger V, Pellissier S, Fauvelle F. A 12-month pilot study outcomes of vagus nerve stimulation in Crohn's disease. Neurogastroenterol Motil. 2020; 32(10): e13911.
14. Yakunina N, Kim SS, Nam EC. Optimization of Transcutaneous Vagus Nerve Stimulation Using Functional MRI. Neuromodulation. 2017; 20(3): 290-300.
15. Parikh A, Lewis G, GholamHosseini H. Home-based respiratory-gated transcutaneous auricular vagus nerve stimulation for rheumatoid arthritis-a feasibility study. Clin Rheumatol. 2026; 45(5): 2627-2638.
16. Sahn B, Pascuma K, Kohn N, Tracey KJ, Markowitz JF. Transcutaneous auricular vagus nerve stimulation attenuates inflammatory bowel disease in children: a proof-of-concept clinical trial. Bioelectron Med. 2023; 9(1): 23.
17. Mao Y, Chen C, Falahpour M. Effects of Sub-threshold Transcutaneous Auricular Vagus Nerve Stimulation on Cingulate Cortex and Insula Resting-state Functional Connectivity. Front Hum Neurosci. 2022; 16: 862443.
18. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. Gut. 2006; 55(6): 749-753.
19. Best WR, Becktel JM, Singleton JW, Kern F Jr. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. Gastroenterology. 1976;70(3): 439-444.
20. Neves ML, Coutinho BD, Speretta GFF Laterality-dependent biological effects of manual acupuncture stimulation of the auricular vagus nerve in women with fibromyalgia: a randomized, sham-controlled trial. Pain Med. 2026; 27(7): 804-813.
21. Pallis AG, Mouzas IA, Vlachonikolis IG. The inflammatory bowel disease questionnaire: a review of its national validation studies. Inflamm Bowel Dis. 2004; 10(3): 261-269.
22. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res. 2002; 52(2): 69-77.
23. Rosas-Ballina M, Ochani M, Parrish WR. Splenic nerve is required for cholinergic antiinflammatory pathway control of TNF in endotoxemia. Proc Natl Acad Sci U S A. 2008; 105(31): 11008-11013.
24. Bonaz B. Is-there a place for vagus nerve stimulation in inflammatory bowel diseases?. Bioelectron Med. 2018; 4:4.
25. Fornaro R, Actis GC, Caviglia GP, Pitoni D, Ribaldone DG. Inflammatory Bowel Disease: Role of Vagus Nerve Stimulation. J Clin Med. 2022; 11(19): 5690.
26. Bonaz B, Sinniger V, Pellissier S. Vagus nerve stimulation: a new promising therapeutic tool in inflammatory bowel disease. J Intern Med. 2017; 282(1): 46-63.
27. Caravaca AS, Levine YA, Drake A, Eberhardson M, Olofsson PS. Vagus Nerve Stimulation Reduces Indomethacin-Induced Small Bowel Inflammation. Front Neurosci. 2022; 15: 730407.
28. Liu CH, Yang MH, Zhang GZ. Neural networks and the anti-inflammatory effect of transcutaneous auricular vagus nerve stimulation in depression. J Neuroinflammation. 2020; 17(1): 54.
29. Burisch J, Zhao M, Odes S. The cost of inflammatory bowel disease in high-income settings: a Lancet Gastroenterology & Hepatology Commission. Lancet Gastroenterol Hepatol. 2023; 8(5): 458-492.
30. Aranow C, Atish-Fregoso Y, Lesser M. Transcutaneous auricular vagus nerve stimulation reduces pain and fatigue in patients with systemic lupus erythematosus: a randomised, double-blind, sham-controlled pilot trial. Ann Rheum Dis. 2021; 80(2): 203-208.
31. Vuyyuru SK, Nguyen TM, Murad MH. Comparative Efficacy of Advanced Therapies for Achieving Endoscopic Outcomes in Crohn's Disease: A Systematic Review and Network Meta-Analysis. Clin Gastroenterol Hepatol. 2024; 22(6): 1190-1199.e15.