ISSN 1881-7831 Online ISSN 1881-784X

DD&T

Drug Discoveries & Therapeutics

Volume 20, Number 2
April 2026

IRCA BSSA

IACMHR

www.ddtjournal.com

DD&T
Drug Discoveries & Therapeutics

ISSN: 1881-7831
Online ISSN: 1881-784X
CODEN: DDTRBX
Issues/Year: 6
Language: English
Publisher: IACMHR Co., Ltd.

Drug Discoveries & Therapeutics is one of a series of peer-reviewed journals of the International Research and Cooperation Association for Bio & Socio-Sciences Advancement (IRCA-BSSA) Group. It is published bimonthly by the International Advancement Center for Medicine & Health Research Co., Ltd. (IACMHR Co., Ltd.) and supported by the IRCA-BSSA.

Drug Discoveries & Therapeutics publishes contributions in all fields of pharmaceutical and therapeutic research such as medicinal chemistry, pharmacology, pharmaceutical analysis, pharmaceutics, pharmaceutical administration, and experimental and clinical studies of effects, mechanisms, or uses of various treatments. Studies in drug-related fields such as biology, biochemistry, physiology, microbiology, and immunology are also within the scope of this journal.

Drug Discoveries & Therapeutics publishes Original Articles, Brief Reports, Reviews, Policy Forum articles, Case Reports, Communications, Editorials, News, and Letters on all aspects of the field of pharmaceutical research. All contributions should seek to promote international collaboration in pharmaceutical science.

Editorial Board

Editorial and Head Office:

Pearl City Koishikawa 603,
2-4-5 Kasuga, Bunkyo-ku,
Tokyo 112-0003, Japan
E-mail: office@ddtjournal.com

Drug Discoveries & Therapeutics

Editorial and Head Office

Pearl City Koishikawa 603, 2-4-5 Kasuga, Bunkyo-ku,
Tokyo 112-0003, Japan

E-mail: office@ddtjournal.com
URL: www.ddtjournal.com

Review

Original Article

Brief Report

Review

DOI: 10.5582/ddt.2026.01020

Artificial intelligence (AI)-assisted ultrasound in clinical trials: Endpoint automation, decentralized monitoring, and regulatory readiness

Kenji Karako[1,*], Jianjun Gao[2]

[1] Department of Surgery, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan;
[2] Department of Pharmacology, School of Pharmacy, Qingdao Medical College, Qingdao University, Qingdao, Shandong, China.

SUMMARY: Ultrasound is among the most widely used imaging modalities in clinical trials, and yet its dependence on operator skill and equipment settings has historically limited the reproducibility of ultrasound-based endpoints in multi-center studies. Artificial intelligence (AI) now addresses this limitation across two complementary dimensions: automated measurement algorithms that quantify cardiac function, organ volume, and vascular parameters with reproducibility approaching or, in select settings, exceeding that of trained human readers and real-time acquisition guidance systems that enable clinicians with no formal sonography training to perform diagnostic-level examinations, making remote and decentralized assessment increasingly feasible. This narrative review synthesizes current evidence and regulatory developments across three interconnected domains. First, use of automated ultrasound to ascertain endpoints has advanced from single-institution validation to prospective and randomized evidence, with deep learning measurement of the left ventricular ejection fraction demonstrating formal equivalence to expert readers across multiple echocardiographic parameters and AI-first workflows shortening the time to diagnosis in a blinded non-inferiority trial. Second, AI-guided ultrasound acquisition by nurses and other non-expert operators has achieved a high rate of diagnostic acceptability in cardiac and pulmonary ultrasound, laying the groundwork for use of ultrasound-based endpoints in decentralized clinical trial designs, as reflected in Food and Drug Administration (FDA) guidance on decentralized trial elements. Third, the regulatory frameworks governing AI-enabled medical devices - including the U.S. FDA's Predetermined Change Control Plan guidance, the EU Artificial Intelligence Act, and internationally harmonized good machine learning practice relevant to Japan and other jurisdictions - increasingly emphasize overlapping principles such as specification of prospective performance, post-marketing oversight, and transparent reporting. Addressing remaining challenges in domain generalization across vendors, subgroup fairness, and algorithm change management during ongoing trials will be essential for AI-assisted ultrasound to fulfill its potential as a robust, scalable endpoint in clinical research worldwide.

Keywords: Artificial intelligence, ultrasound, clinical trials, decentralized clinical trials, automated endpoint measurement

1. Introduction

Ultrasound (US) imaging occupies a distinctive position as a clinical trial methodology. Unlike computed tomography or magnetic resonance imaging, US is non-ionizing, the equipment is portable, and scans can be acquired in real time, enabling its use across a spectrum of trial settings—from centralized imaging facilities in pivotal phase III studies to point-of-care assessments in resource-limited environments (*1*). Cardiac function endpoints such as the left ventricular ejection fraction (LVEF), as well as measurements of organ volume and vascular parameters, are routinely incorporated as primary or secondary endpoints in cardiovascular, oncological, and obstetric trials (*1*). The fact that clinical research is being performed around the world in settings with fewer resources, accelerated by disruptions to conventional trial infrastructure due to the COVID-19 pandemic, has further heightened the demand for imaging modalities that can function reliably outside tertiary hospitals.

Nonetheless, ultrasound has historically been the imaging modality most susceptible to measurement variability. Image quality and diagnostic accuracy depend heavily on operator skill, probe selection, patient positioning, and equipment settings—factors that vary

across sites, sonographers, and time points within the same trial. LVEF estimation using conventional two-dimensional echocardiography involves significant inter-observer variability. 2D methods result in minimal detectable differences of approximately 10–13%, a magnitude that may obscure meaningful treatment-related changes in cardiovascular outcome trials (*2*). Domain shift—the degradation of performance when a model is applied to data from a different institution, scanner, or patient population—is a further intrinsic challenge in ultrasound that constrains the external validity of both human readers and automated algorithms (*3,5*). These characteristics have limited the scalability of ultrasound-based endpoints in multi-center studies, where standardization of acquisition and interpretation is essential for valid cross-site comparisons.

Artificial intelligence (AI), and deep learning in particular, has emerged as a transformative tool to overcome these limitations. Convolutional neural networks (CNNs), vision transformers (ViTs), and self-supervised learning (SSL) frameworks have demonstrated the ability to automate ultrasound image analysis with reproducibility meeting or exceeding that of human experts when performing standardized tasks (*3,5,8,9*). A landmark video-based model for LVEF estimation had a mean absolute error (MAE) of 4.1% when using internal data and 6.0% when using an independent external dataset, with comparable inter-clinician variability, thus laying the foundation for automated echocardiography as an endpoint (*3*). A blinded, randomized non-inferiority trial that evaluated an AI-first echocardiographic workflow noted significantly lower rates of clinically meaningful LVEF discordance (> 5%) between initial assessment and the final cardiologist reading in the AI arm versus the sonographer arm (16.8% *vs.* 27.2%; $p < 0.001$ for superiority). This established a precedent for use of AI-assisted measurement as a primary trial endpoint in randomized controlled trials (RCTs) (*4*). In terms of ultrasound acquisition, AI-guided real-time probe-positioning systems have enabled clinicians with no prior sonography training—as well as nurses—to perform diagnostic-quality 10-view echocardiography, with 100% of primary-endpoint examinations meeting diagnostic acceptability criteria across all four primary endpoints in a prospective international noninferiority study (*6*). These advances in cardiac imaging suggest a shift in the central question from whether AI can assist ultrasound to how such assistance can be validated, standardized, and regulated within the formal framework of a clinical trial.

Three interrelated challenges define the current frontier. First, use of automated measurement algorithms to reliably ascertain clinical trial endpoints and whether they meet standards for accuracy, reproducibility, and external validity across heterogeneous scanning environments and patient populations encountered in multi-center studies must be validated. Second, AI-guided ultrasound acquisition foreshadows the potential for decentralized clinical trials (DCTs)—designs in which some or all imaging assessments are performed at local facilities or remotely rather than at designated imaging centers—but raises new questions about quality assurance, operator training, and data integrity. A regulatory framework for such designs was finalized by the U.S. Food and Drug Administration (FDA) in September 2024 (*7*). Third, the regulatory frameworks governing AI-based medical devices are evolving rapidly across major jurisdictions, and trial sponsors must navigate these requirements when incorporating AI-assisted ultrasound as an endpoint or trial tool.

This review addresses all three challenges systematically. We first examine the evidence base for using automated ultrasound in endpoint measurement, covering technical approaches, clinical validation studies up to and including RCTs, and the implications that a domain shift would have in multi-center deployment. We then consider AI-guided image acquisition in the context of decentralized trials and emergency setting trials, synthesizing prospective clinical evidence and identifying barriers to widespread implementation. Finally, we review the current regulatory landscape across the United States, Japan, and the European Union and outline the methodological standards—reporting guidelines, evaluation frameworks, and privacy-preserving learning architectures—that trial designers should incorporate to ensure reproducibility and regulatory compliance. Together, these three components represent the infrastructure required for AI-assisted ultrasound to fulfill its potential as a robust clinical trial tool. The conceptual relationship among these components is illustrated in Figure 1. The literature search strategy and inclusion criteria underlying this synthesis are described in the Methods section below.

2. Methods

This narrative review synthesizes recent literature on AI-assisted ultrasound in clinical trials. Relevant studies were identified through searches of PubMed, Embase, IEEE Xplore, and the Cochrane Library using combinations of the terms "ultrasound," "echocardiography," "artificial intelligence," "deep learning," "clinical trial," and "decentralized clinical trial." Additional sources were identified in references listed in key articles and from the websites of major regulatory agencies, including the FDA, Pharmaceuticals and Medical Devices Agency (PMDA)/Ministry of Health, Labor, and Welfare (MHLW), International Medical Device Regulators Forum (IMDRF), and the EU Official Journal.

Priority was given to prospective studies, randomized or non-inferiority trials, formal external validation studies, and primary regulatory guidance documents published up to March 2026. When preprints were cited,

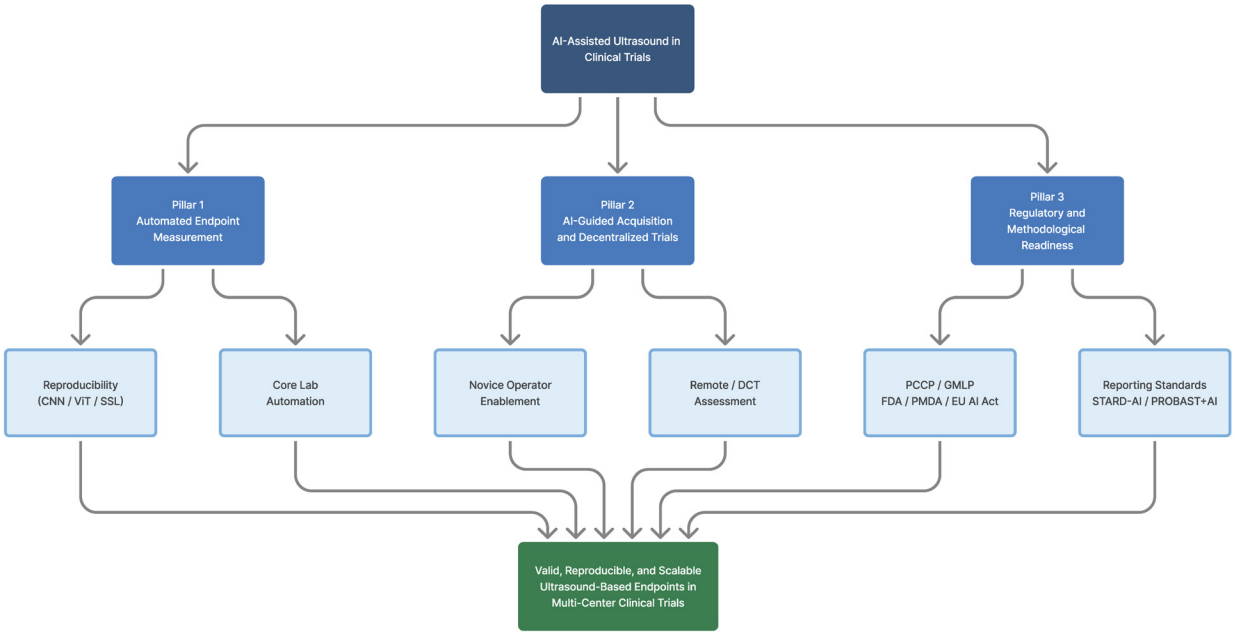

Figure 1. Conceptual framework of AI-assisted ultrasound in clinical trials. Three interdependent pillars are required for AI-assisted ultrasound to function as a reliable clinical trial tool. Pillar 1 (automated endpoint measurement) and Pillar 2 (AI-guided ultrasound acquisition and decentralized trials) each generate the evidence and infrastructure needed for valid endpoints. Pillar 3 (regulatory and methodological readiness) provides the governance layer that enables deployment across sites and regulatory acceptance. All three pillars converge on the goal of valid, reproducible, and scalable ultrasound-based endpoints in multi-center trials.

this was made explicit in the references and interpreted as emerging rather than definitive evidence. This was a narrative synthesis rather than a systematic review, so formal study selection, risk-of-bias adjudication, and meta-analytic pooling procedures were not performed.

3. Using automated ultrasound to ascertain clinical trial endpoints

3.1. Technological foundations

Three families of deep learning architecture underpin automated ultrasound measurement: CNNs, ViTs, and SSL frameworks. Each addresses a different constraint inherent to clinical ultrasound data.

CNNs were the first deep learning approach to ultrasound that was clinically validated. Their inductive biases—local feature detection and translational invariance—suit the spatially structured patterns of B-mode images, and their computational efficiency enables near-real-time inference on consumer hardware. The most consequential demonstration of CNN-based ultrasound analysis in a clinical trial context is EchoNet-Dynamic, which applied a video-based convolutional neural network to full echocardiographic video sequences to estimate the LVEF and segment the left ventricle (*3*). The model achieved an MAE of 4.1% on the internal held-out test set, which is well within the range of inter-observer variability among clinicians, and generalized to an independent external healthcare system with an MAE of 6.0%, demonstrating robust external validity for a video-based AI echocardiography system (*3*).

ViTs address a limitation of CNNs: their localized receptive fields hamper the integration of global spatial context, which is important for tasks such as view classification and whole-image quality assessment. By treating image patches as sequential tokens and applying self-attention mechanisms, ViTs capture long-range dependencies across the field of view. A systematic review of 69 publications covering ViT applications across cardiac, breast, prostate, and fetal ultrasound highlighted their strength in capturing global contextual information. At the same time, multiple studies reported that hybrid architectures combining convolutional feature extraction with transformer attention can improve performance, and particularly in settings with limited training data (*8*).

SSL exploits the large volumes of unlabeled ultrasound images available in clinical archives—data that are impractical to annotate exhaustively—by training models to reconstruct, contrast, or predict features of their own input without manual labels. This approach is especially valuable for reducing dependence on expert annotation in specialized domains. At the foundation model scale, UltraFedFM pre-trained a shared model *via* federated learning across 16 institutions in 9 countries using over one million unlabeled ultrasound images, achieving an area under the receiver operating characteristic curve (AUROC) of 0.927 for disease classification and a Dice coefficient of 0.878 for lesion segmentation across downstream tasks (*9*). The implications for multi-center trial infrastructure

are discussed further in Section 5. The principal characteristics, ultrasound-specific limitations, and representative evidence for each of the three families of architecture are compared in Table 1.

Cutting across all three families of architecture is the emerging field of explainable AI (XAI), which generates human-interpretable rationales for model predictions—for example, attention heatmaps highlighting the myocardial regions driving an LVEF estimate, or saliency maps indicating which image zones triggered a lesion classification. XAI is not an alternative architecture but a post-hoc or intrinsic interpretability layer that can be applied to CNN, ViT, and SSL-derived models alike. Its clinical trial relevance is twofold: regulators increasingly expect AI-enabled medical devices to provide interpretable outputs sufficient for clinical oversight, and FUTURE-AI's explainability dimension recommends that the basis for an AI decision be communicated to the clinical reviewer in a way that supports, rather than replaces, human judgment (*17*). For trial endpoints specifically, XAI outputs can serve as audit evidence during regulatory review of discordant AI-human measurements and as a mechanism for detecting systematic model failure modes—such as dependence on image artifacts rather than genuine anatomical features—before adoption.

3.2. Clinical validation of using AI to ascertain trial endpoints

Converting an AI measurement algorithm from conjecture into a validated clinical trial tool requires a structured hierarchy of evidence. This section first presents a pragmatic five-level framework synthesized from the literature on AI-assisted echocardiography—where the evidence base is currently most complete—and then considers how far analogous evidence has emerged in non-cardiac areas. Representative studies are mapped to each level in Figure 2.

Level 1—Internal validation using held-out data. The minimum requirement for any AI measurement system is that its performance has been determined using a held-out test set from the same institution or dataset as training data. When estimating the LVEF, EchoNet-Dynamic achieved an MAE of 4.1% with its internal held-out set—approaching the inter-clinician variability for the LVEF reported in the same clinical environment—establishing a minimum credibility threshold (*3*).

Level 2—External validity using independent data. In addition, EchoNet-Dynamic's performance in an independent external healthcare system with different scanning environments yielded an MAE of 6.0%—

Table 1. The three families of architecture in terms of typical endpoints, key limitations in the ultrasound context, and representative references

Architecture	Typical endpoints	Ultrasound-specific limitations	Representative reference
CNN (2D/3D)	LVEF, segmentation, view classification	Device–vendor domain shift; limited global context	Ouyang *et al*., Nature 2020 (*3*)
Vision Transformer	Classification, segmentation, quality grading	Requires larger training sets; sensitive to domain shift	Vafaeezadeh *et al*., Diagnostics 2024 (*8*)
SSL/Foundation model	Multi-task pre-training → downstream fine-tuning	Pre-training dataset diversity is a bottleneck	Jiang *et al*., NPJ Digit Med 2025 (*9*)

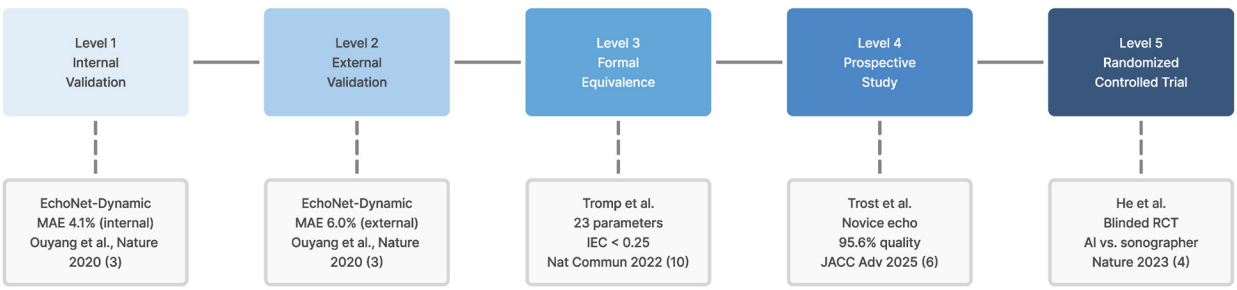

Figure 2. Hierarchy of clinical evidence for use of AI-assisted ultrasound to ascertain endpoints. Five levels of evidence for use of AI-assisted ultrasound to ascertain clinical trial endpoints. They are shown here, along with landmark studies in AI-assisted echocardiography. Level 1 involves establishing internal performance based on held-out data from the institution developing the AI; Level 2 involves demonstrating external validity using independent data from a separate institution or scanning environment; Level 3 involves applying a pre-specified equivalence criterion (*e.g.*, the individual equivalence coefficient) to repeated human expert assessments; Level 4 involves prospectively evaluating ultrasound performed by non-expert operators under trial conditions; and Level 5 indicates the steps in a randomized, blinded evaluation. A trial sponsor should pre-specify the minimum level of evidence required for endpoint adoption in the trial protocol; Level 3 is the recommended minimum for primary endpoints.

within the range of human inter-clinician variability (*3*). This constitutes the first large-scale demonstration of cross-site external validity for a video-based AI model for echocardiography and establishes the principle that cross-site validation should be prospectively planned rather than retrospective. An algorithm remains at Level 1 until this external step has been completed.

Level 3—Formal equivalence validation. Tromp *et al.* applied a more rigorous standard, comparing deep learning measurements of 23 echocardiographic parameters—including the ejection fraction, ventricular volume, and Doppler-derived filling pressure—from 600 participants to three independent repeated measurements by trained sonographers (*10*). The primary criterion was the individual equivalence coefficient, a statistic that quantifies whether algorithm-human disagreement is smaller than human-human disagreement. For all 23 parameters, the upper bound of the 95% confidence interval (CI) for the individual equivalence coefficient fell below the pre-specified non-inferiority threshold of 0.25, indicating that the deep learning workflow was at least as reproducible as expert human assessment (*10*). This study design, which pre-registers the equivalence criterion before data analysis, provides the most relevant methodological template for regulatory and endpoint validation purposes.

Level 4—Prospective evaluation in the intended operator population. An algorithm validated at Level 3 has been assessed in a controlled laboratory setting but not necessarily in the hands of the operator population that will use it in a trial. Trost *et al.* addressed this gap in a two-site international study (240 participants) in which nurses—the intended DCT operator—used an AI-guided system to perform echocardiography, achieving 100% diagnostic acceptability for all four pre-specified primary endpoints (LV size, global LV function, RV size, and pericardial effusion), identical to the expert reference group (*6*). This level of evidence is analogous to a phase II clinical study for a drug: it confirms that a validated algorithm remains fit-for-purpose when used by non-expert operators under real trial conditions. That study is discussed in detail in Section 4.2.

Level 5—Randomized controlled trial of the AI-enhanced workflow. He *et al.* conducted a blinded, randomized non-inferiority trial at Cedars-Sinai Medical Center comparing AI-assisted initial LVEF assessment to standard sonographer assessment (*4*). The pre-specified primary endpoint was the proportion of studies in which the initial assessment (AI or sonographer) differed from the reviewing cardiologist's final LVEF reading by more than 5%. Among 3,495 studies, this discordance rate was significantly lower in the AI arm (16.8%) than in the sonographer arm (27.2%; difference −10.4%, 95% CI −13.2 to −7.7; $p < 0.001$ for superiority), and blinding was successfully maintained (blinding index: 0.088) (*4*). This trial demonstrated that an AI-first echocardiographic workflow can be evaluated within the formal infrastructure of a randomized, blinded clinical study—a critical precedent for trials seeking to use AI measurement as their primary endpoint.

3.2.1. Generalizability to non-cardiac ultrasound domains

The five-level hierarchy described above is most clearly evident in AI-assisted echocardiography, where the sequence from internal validation to RCT evidence now spans multiple independent groups. Evidence from non-cardiac domains is accumulating, though most applications remain at Levels 1–3. A prospective two-country diagnostic accuracy study of fetal ultrasound demonstrated that deep learning estimation of gestational age from blind ultrasound sweeps achieved an MAE of 3.2 days. The AI system met a pre-specified equivalence margin (± 2 days) relative to standard sonographer-performed biometry when both were evaluated with first-trimester crown–rump length as the reference standard (*26*). A complementary four-center prospective study conducted across Moroccan healthcare centers developed and prospectively evaluated a deep learning system for end-to-end automation of fetal biometry and amniotic fluid assessment from ultrasound cine-loops. The model achieved limits of agreement with expert measurements that were narrower than reported intra- and inter-observer variability among sonographers, indicating reproducibility comparable to or better than human performance (*27*). A six-center prospective multicenter study of hepatic ultrasound developed and internally validated a deep learning model for multi-stage liver fibrosis classification (S0–S4) using paired high-frequency ultrasound images of the liver and spleen, with histologic staging as the reference standard. The model achieved an AUROC of 0.87–0.94 across clinically relevant fibrosis thresholds and demonstrated the potential to reduce biopsy rates in simulated clinical pathways (*28*). The Controlled ARTHUR Trial of musculoskeletal ultrasound prospectively evaluated a robotic ultrasonography system combined with AI-based synovitis scoring in rheumatoid arthritis, comparing its performance to that of expert-performed ultrasound. The study noted moderate agreement between robotic and human assessments (intraclass correlation coefficients of 0.59–0.64) and limited diagnostic accuracy (approximately 59%), highlighting both the feasibility and current limitations of operator-independent ultrasound acquisition and automated scoring (*29*). Across these domains, emerging evidence suggests that AI-based measurement can achieve levels of reproducibility comparable to expert human assessment in selected applications, while prospective multi-center studies are beginning to establish the Level 2–3 evidence base required for further validation for clinical and trial use.

3.3. Challenges in multi-center deployment

Despite this evidence, three challenges constrain the use of automated ultrasound to ascertain endpoints across heterogeneous multi-center trial networks.

Domain shift—performance degradation due to differences in the scanner vendor, probe frequency, imaging protocol, or patient population—is the most consistently cited obstacle. Recently, Universal UltraSound Image Challenge 2025 evaluated general-purpose ultrasound AI models' performance on data aggregated from public datasets and partner hospitals across multiple anatomical regions, and the competition revealed measurable degradation of performance on data from institutions not represented in the training dataset (*11*). Although this evidence is still preliminary, it reinforces an important operational point: domain generalization must be explicitly validated, not assumed, before multi-center deployment.

Subgroup performance gaps are a related concern. Differences in AI measurement accuracy across patient subgroups—potentially defined by sex, age, body mass index, or scanner type—may introduce systematic bias that inflates or attenuates between-arm differences in a trial with imbalanced enrollment, analogous to differential misclassification in conventional measurement. Sex-related performance gaps have been empirically documented in musculoskeletal ultrasound AI (*12*); analogous biases are plausible across cardiac, hepatic, and vascular domains, where training dataset composition is similarly skewed. Pre-registration of analyses of subgroup accuracy and reporting of subgroup-stratified performance metrics are therefore essential components of validating use of AI to ascertain endpoints.

Public data availability remains a bottleneck for algorithm development and replication. A systematic catalog of publicly available ultrasound resources identified 72 datasets and 56 open-source models as of September 2025, with notable underrepresentation of fetal (5 datasets) and prostate (4 datasets) content relative to the frequency of their clinical use (*13*). For investigators seeking to independently validate algorithms before adoption in a trial, this scarcity of benchmark data is a practical impediment.

3.4. Reporting standards

Transparent reporting is a prerequisite for regulators, site monitors, and independent reviewers to evaluate whether an AI endpoint meets trial quality standards. Two complementary guidelines are particularly relevant. STARD-AI, published in Nature Medicine in 2025, extends the established STARD checklist with AI-specific reporting items covering dataset composition, model evaluation, and bias or fairness considerations (*14*). PROBAST+AI, published in the British Medical Journal in 2025, provides a structured tool for assessing risk of bias and applicability in predictive model studies using regression or AI (*15*). These frameworks do not themselves constitute regulatory requirements, but they increasingly shape the evidence packages that regulators, sponsors, and reviewers expect when proposing use of an AI system in a trial. The alignment between common trial study designs and their most relevant reporting or evaluation framework is outlined in Table 2.

4. AI-guided ultrasound acquisition and decentralized ultrasound monitoring

4.1. Mechanisms of AI-guided ultrasound acquisition and their significance for clinical trials

Automated measurement algorithms, however accurate, are only as reliable as the images on which they operate. Suboptimal image acquisition—incorrect probe orientation, inadequate acoustic coupling, or failure to obtain the prescribed view—introduces a source of variability that no downstream algorithm can fully correct. AI-guided ultrasound addresses this upstream constraint by providing real-time, computer-generated feedback to the operator during image acquisition, decoupling diagnostic quality from individual sonographer experience.

The technical architecture of guided image acquisition systems typically combines two deep learning components: a view classification network that identifies which of the target views has been achieved, and a guidance-providing module that translates the current probe position into actionable instructions—expressed as directional cues on a display or auditory prompts—to direct the operator toward the correct plane. These components operate on live video input with inference latencies low enough for interactive use in clinical settings in real time. The classification network may be a CNN, a ViT, or a hybrid architecture; the critical design requirement is the ability to distinguish near-correct views from target views with sufficient specificity to avoid issuing conflicting instructions, which is particularly challenging in echocardiography given the anatomical proximity of standard imaging planes.

Table 2. Trial study designs within the most relevant reporting or evaluation frameworks

Study type	Primary framework	Key AI-specific requirements
Diagnostic accuracy study	STARD-AI (*14*)	Dataset provenance, subgroup reporting, bias/fairness
Predictive model development/validation	PROBAST+AI (*15*)	Risk of bias across participants, predictors, outcomes, analysis
Early-stage real-world evaluation	DECIDE-AI (*16*)	Human–AI interaction, safety monitoring
Comprehensive evaluation of implementation	FUTURE-AI (*17*)	Fairness, universality, traceability, robustness, explainability

When conducting a clinical trial, the significance of reliable AI-guided ultrasound acquisition extends beyond quality control into the trial design itself. If novice operators—nurses, first-responders, or community health workers without formal sonography training—can obtain diagnostic-quality ultrasound images with AI guidance, then imaging assessments that would previously have required a patient to travel to an imaging facility can instead be performed at local clinics, remote research sites, or in participants' homes. This capability is a prerequisite for a DCT involving imaging-dependent endpoints and directly addresses the operational constraints that have historically limited ultrasound-based endpoints to site-intensive trial designs.

4.2. Prospective clinical evidence

The clinical evidence base for AI-guided ultrasound acquisition has expanded substantially in the past two years and now spans cardiac, pulmonary, vascular, obstetric, and abdominal applications.

Cardiac ultrasound. The most rigorous assessment to date is a prospective, international, noninferiority study conducted at two sites (United States and France) by Trost *et al*. (*6*). Nurses without prior echocardiography experience used an AI-guided system to acquire 10 standard transthoracic echocardiographic views of 240 patients over a nine-month period (November 2023 to August 2024). For all four pre-specified primary endpoints—assessment of left ventricular size, global left ventricular function, right ventricular size, and nontrivial pericardial effusion—100% of novice-performed examinations met diagnostic acceptability criteria, with no difference compared to expert-performed examinations. Examinations were independently reviewed by five cardiologists blinded to operator identity, and strong agreement was observed between AI-guided image acquisition by novices and that by experts across multiple parameters. Of methodological note, the study prospectively defined its performance thresholds and employed blinded expert adjudication, providing high-quality prospective evidence supporting the feasibility of AI-guided ultrasound by non-expert operators in clinical settings.

Pulmonary ultrasound. Baloescu *et al*. extended the acquisition-guidance paradigm to pulmonary ultrasound in a multicenter study spanning four sites (*18*). Non-expert operators used the system to perform a standardized eight-zone protocol in patients with acute respiratory symptoms. Diagnostic-quality images were obtained in 98.3% of examinations, with no significant difference compared to expert-performed examinations. Importantly, the study focused on acquisition quality rather than diagnostic accuracy, demonstrating that AI guidance enables non-expert operators to reliably obtain images suitable for clinical interpretation. These findings are relevant to decentralized trial settings, where imaging must be performed by a distributed workforce without extensive ultrasound expertise.

Vascular ultrasound. Speranza *et al.* evaluated AI guidance in diagnosis of deep vein thrombosis in a setting where non-expert operators performed compression ultrasound of the lower extremities (*19*). In a multicenter study of 381 patients across 11 UK hospitals, scans acquired with AI guidance were reviewed by qualified clinicians and compared to standard examinations by imaging specialists. Emergency medicine reviewers achieved a sensitivity of 95-98% and a specificity of 97-100%. Importantly, the diagnostic workflow incorporated clinician review, with indeterminate or insufficient-quality scans requiring further assessment. This study highlights that AI-guided ultrasound can be integrated with expert review to maintain diagnostic performance while potentially reducing the need for universal expert-performed examinations.

Obstetric ultrasound. The acquisition-guidance paradigm extends to obstetric ultrasound, where access to trained sonographers is a primary constraint on the feasibility of maternal-fetal medicine trials in limited-resource settings. The prospective two-center study by Stringer *et al.* described in Section 3.2.1—in which novice operators with minimal training in Zambia performed freehand blind sweeps using an AI-guided portable device—yielded gestational age estimation that was non-inferior to that of credentialed sonographers, confirming that AI guidance can substitute for formal training with an obstetric endpoint that is image acquisition-dependent (*26*). Gomes *et al.* validated a complementary mobile-optimized AI system capable of running without Internet connectivity on a handheld device, designed specifically for novice operator use in resource-limited settings (*31*). The system provided feedback scores reflecting sweep quality, which can assist operators in identifying inadequate acquisitions, and achieved gestational age estimation to within -1.4 ± 4.5 days of standard sonographer biometry; fetal malpresentation—a clinically important finding that would qualify as a secondary endpoint in preterm birth prevention trials—was detected with an AUROC of 0.977 (95% CI 0.949–1.000) (*31*). Together these studies demonstrate that operator-independent obstetric imaging protocols are feasible for endpoints central to global reproductive health trials, including standardization of gestational age and screening for malpresentation.

Abdominal ultrasound. Chiu *et al*. conducted a prospective evaluation of AI-guided abdominal aortic aneurysm screening in which 10 nurses with no prior ultrasonography experience each performed fifteen scans using a real-time deep learning detection algorithm (*32*). Diagnostic image quality was achieved in 87.5% of nurse-performed scans, compared to 91.3% for physician-performed scans ($p = 0.310$), establishing equivalence at a clinically acceptable threshold. Abdominal aortic aneurysm detection sensitivity was

100% and specificity was 97.8% (AUROC 0.975), with an MAE for measurement of the aortic diameter of 2.8 mm—within the threshold conventionally used for surveillance decisions in screening programs. These findings suggest that AI-guided ultrasound acquisition may enable non-expert operators to perform abdominal ultrasound screening with performance comparable to experienced physicians, supporting its potential applicability in scalable screening and decentralized trial settings.

4.3. DCTs: Regulatory framework and implementation requirements

The FDA's September 2024 final guidance on conducting clinical trials with decentralized elements defines DCTs as studies in which trial-related activities occur outside traditional clinical trial sites, including through local healthcare providers, in-home visits, or telehealth and digital health technologies (*7*). The guidance recognizes that imaging and other clinical assessments may be performed at local facilities and emphasizes the need for standardized protocols, appropriate training of personnel, and robust data quality and integrity monitoring. Sponsors are also expected to specify data origin, data flow, and monitoring procedures within the trial protocol.

For ultrasound-based imaging endpoints specifically, the FDA's earlier guidance on Clinical Trial Imaging Endpoint Process Standards, finalized in 2018, outlines recommended process standards applicable across clinical trial designs (*20*). These include pre-specified imaging acquisition protocols, operator qualification and training, equipment standardization and quality control procedures, as well as requirements for image transfer, storage, and centralized interpretation where appropriate (*20*). Sponsors incorporating AI-guided ultrasound acquisition into decentralized trial designs should ensure that image quality and protocol adherence are maintained when performed by the intended user population.

Despite regulatory clarity and mounting clinical evidence, substantial barriers to implementation remain. The COMPASS-AI survey, which collected structured responses from 1,154 healthcare professionals across multiple countries, found that 81.1% expressed enthusiasm for AI-assisted POCUS, and yet identified training and education (27.1% of respondents) and clinical validation and evidence (17.5%) as the two most frequent barriers to adoption (*21*). Additional concerns included limitations in infrastructure, workflow integration, lack of standardized guidelines, and uncertainty regarding liability. These findings suggest that, in DCT settings, institutional readiness may represent a critical limiting factor.

4.4. Emergency and resource-constrained settings

The evidence reviewed in Sections 4.2 and 4.3 addresses the operational challenges of decentralizing clinical trials within established healthcare systems; the same AI-guided ultrasound acquisition capabilities also extend the scope of the evidence reviewed in Section 4.2 into emergency and resource-constrained settings, where conventional sonographer-dependent protocols are entirely infeasible and where endpoint integrity would otherwise be unachievable.

The COVID-19 pandemic provided a real-world stress test of decentralized imaging workflows. Widespread adoption of POCUS for cardiopulmonary triage in emergency departments and temporary care settings—where radiological infrastructure was limited and trained sonographers were scarce—demonstrated that portable ultrasound can be used effectively with simplified, protocolized acquisition approaches under emergency conditions. Although primarily used for clinical decision-making rather than formal trial endpoints, these experiences suggest that structured and protocolized ultrasound workflows can be implemented with practical feasibility in resource-constrained environments. This has important implications for decentralized trial design: AI-guided ultrasound acquisition systems may further enhance standardization and reproducibility, particularly when operator expertise and participant mobility are limited.

In low- and middle-income countries (LMICs), endpoint validity carries an additional dimension: the underrepresentation of LMIC populations in clinical trials is a recognized threat to the external validity and generalizability of trial results. A scoping review of 29 studies on AI-enabled POCUS applications in limited-resource settings identified pulmonology medicine and obstetrics as the most frequently investigated domains, with many studies focusing on diagnostic support and workflow facilitation (*22*). Implementation constraints in LMICs differ from those in high-income countries, with challenges including infrastructure limitations, lack of standardization, and limited data availability rather than solely operator training. The fetal ultrasound validation data from Zambia (*26*) and Morocco (*27*) reviewed in Section 3.2.1 provide direct evidence that AI-enabled systems can achieve performance comparable to expert standards in limited-resource environments. While the Zambian study demonstrated the equivalence of AI-assisted image acquisition by novice operators, the Moroccan study showed that automated measurement pipelines can reach human-level reproducibility. Together, those studies support the feasibility of scalable trial endpoints in LMICs.

5. Regulatory and methodological readiness

5.1. Evolving regulatory frameworks for AI-enabled medical devices

AI-based medical devices, including those with

automated ultrasound or guided image acquisition, are subject to regulatory oversight in all major jurisdictions. The regulatory treatment of these devices is undergoing rapid change as authorities develop frameworks that accommodate the unique characteristics of AI: continuous learning from post-deployment data, performance that may shift over time, and decision processes that resist traditional static software validation. An overview of the key regulatory features across all three jurisdictions is provided in Table 3.

United States. The FDA regulates AI-enabled ultrasound devices as Software as a Medical Device (SaMD) under the device provisions of the Federal Food, Drug, and Cosmetic Act. As of March 2026, the FDA had authorized more than 1,400 AI-enabled medical devices (*33*). The central regulatory challenge for AI-assisted endpoints is managing modifications: unlike conventional software, AI/ML models may require iterative updates as new training data become available, and each modification that could affect safety or effectiveness may trigger additional regulatory review under traditional 510(k) or De Novo pathways. The Predetermined Change Control Plan (PCCP) framework, for which the FDA finalized guidance in December 2024, addresses this challenge by allowing sponsors to pre-specify the categories of changes they anticipate making to an algorithm—such as retraining on additional data or refining post-processing thresholds—and to obtain prospective FDA agreement that such changes may be implemented without a new marketing submission, provided performance monitoring confirms they remain within pre-specified bounds (*23*). The Good Machine Learning Practice (GMLP) guiding principles, first articulated jointly by the FDA, Health Canada, and the UK's Medicines and Healthcare products Regulatory Agency and later elaborated through IMDRF, provide upstream development expectations applicable across AI/ML medical devices regardless of pathway (*25,30*). These principles cover data management, reference standard selection, and specification of clinically meaningful performance targets, and they complement the PCCP by defining the quality baseline against which predetermined changes should be evaluated.

Japan. Japan's regulatory approach to AI-enabled medical devices has developed alongside international SaMD and GMLP frameworks while retaining its own approval and consultation pathways. The original GMLP guiding principles were published jointly by the U.S. FDA, Health Canada, and the UK's Medicines and Healthcare products Regulatory Agency in 2021 (*25*) and were subsequently reflected in IMDRF guidance, in whose activities Japan participates (*30*). In Japan, AI-enabled medical devices are reviewed within the existing medical device regulatory framework administered by the MHLW and PMDA (*34*), and PMDA has established dedicated review and consultation pathways for SaMD (*35*). Risk management considerations are informed by internationally recognized medical device standards, including ISO 14971, alongside SaMD concepts developed through IMDRF. For qualifying innovative products, the SAKIGAKE Designation System may offer priority consultation and review for devices addressing unmet clinical needs (*36*).

European Union. The EU Artificial Intelligence Act (Regulation (EU) 2024/1689), which entered into force on August 1, 2024, applies on a phased timetable (*24*). The Act becomes generally applicable on August 2, 2026, while high-risk AI systems embedded in regulated products such as medical devices have an extended transition period until August 2, 2027 (*24*). AI-enabled ultrasound systems that guide image acquisition or automate measurement will usually need to satisfy the AI Act along with the Medical Device Regulation (MDR) or In Vitro Diagnostic Regulation (IVDR), rather than instead of them. This means sponsors and developers should anticipate requirements for technical documentation, risk management, human oversight, and post-market monitoring that are complementary to CE-marking obligations under product law. The phased implementation timeline gives developers a defined window to adapt existing systems, but it also means that regulatory planning for EU-inclusive trials should begin well before the 2027 deadline.

5.2. Implications for clinical trial design

The regulatory frameworks described above have direct implications for how AI-assisted ultrasound will be used to ascertain endpoints throughout the duration of a clinical trial.

PCCP and intra-trial model changes. A critical but underappreciated issue in trials using AI to ascertain

Table 3. The key regulatory features across the three jurisdictions

Jurisdiction	Key framework	Change management	Data requirements
United States (FDA)	SaMD / PCCP (*23*)	Predetermined changes pre-agreed with the FDA	GMLP principles (*25*)
Japan (PMDA)	SaMD / ISO 14971 (*34,35*)	Variation submission; IMDRF GMLP alignment (*30*)	Pseudonymized data guidelines (MHLW)
European Union	EU AI Act (*24*) + MDR/IVDR	Post-market monitoring + PCCP equivalent	Technical documentation per AI Act Annex IV

endpoints is algorithm stability. If a trial spans two or more years—as is typical for cardiovascular outcome studies—the sponsor may wish to retrain or update the algorithm on expanded data during the trial period to correct drift or improve accuracy in newly enrolled subpopulations. Under conventional validation paradigms, any modification would require re-validation of the endpoint definition and potentially revision of the statistical analysis plan. The PCCP framework provides a mechanism for managing such changes in a pre-specified, regulatory-compliant manner: if the scope and performance bounds of anticipated modifications are agreed upon with the regulator before trial initiation, updates can proceed without undermining endpoint integrity (*23*). Trial sponsors should therefore consider developing a PCCP for endpoints ascertained using AI in parallel with protocol development, treating algorithm change management as an essential element of defining endpoints.

Post-marketing performance monitoring. FDA requirements and the emerging EU AI Act compliance framework both point toward ongoing performance monitoring of AI systems while in use. For trial sponsors using AI-assisted endpoints, this translates into continued collection of performance data—often through a random or risk-based sample of AI measurements compared against a human reference—throughout the study. Integrating this monitoring into the workflow of the central imaging laboratory adds operational complexity but can provide early warning if systematic performance degradation is detected.

Statistical analysis plan (SAP) considerations for AI endpoints. A frequently overlooked dimension of incorporating AI measurements into clinical trials is their pre-specification in the SAP. Unlike conventional imaging endpoints, AI measurement algorithms produce point estimates that may be accompanied by model confidence scores or prediction intervals; the SAP must specify whether these uncertainty outputs will be used as covariates, to exclude low-confidence measurements from the primary analysis, or to trigger adjudication by a human reader. The equivalence or non-inferiority margin for the AI endpoint should be determined in consultation with regulators before database lock and should be stated explicitly in the SAP rather than inferred post-hoc. When the AI measurement serves as a surrogate endpoint, the statistical linkage between the surrogate and the clinical outcome of interest should be defined and cited in the protocol, following the Prentice criteria or an equivalent framework for surrogate endpoint validation. Finally, subgroup performance analyses—stratified by sex, age, body mass index, and scanner type—should be pre-specified as secondary endpoints rather than exploratory analyses, given documented evidence of differential AI performance in musculoskeletal ultrasound (*12*) and the recognized risk of analogous biases across imaging domains.

5.3. Federated learning as a regulatory-compliant infrastructure

The four reporting and evaluation frameworks described in Section 3.4—STARD-AI, PROBAST+AI, DECIDE-AI, and FUTURE-AI—have important implications for regulatory preparedness even though they are not, by themselves, binding regulatory instruments. Applied prospectively rather than retrospectively, they help shape the pre-approval evidence package and the post-deployment monitoring structure that regulators and sponsors increasingly expect. FUTURE-AI's emphasis on traceability is especially relevant: algorithm version, training data provenance, and inference settings should be logged for endpoint-generating systems so that discrepancies can be audited if they arise during trial oversight or regulatory review.

Federated learning architectures provide a technical mechanism through which PCCP-defined model updates may be implemented across sites, rather than directly fulfilling the PCCP requirements themselves. By enabling each trial site to compute local model updates on its own patient data and contribute only aggregated gradient information to a central server, federated learning allows iterative model refinement without centralizing identifiable patient records, substantially reducing cross-site data transfer risks and achieving compliance with the General Data Protection Regulation and Act on the Protection of Personal Information data minimization principles. Complementary privacy-preserving measures—including differential privacy and secure aggregation protocols—are recommended to guard against gradient inversion attacks that can partially reconstruct training data from shared gradients. UltraFedFM demonstrated this approach at scale, pre-training across 16 institutions in nine countries using over one million unlabeled images to achieve an AUROC of 0.927 for disease classification and a Dice coefficient of 0.878 for lesion segmentation (*9*). For a multi-center trial using a PCCP-governed AI endpoint, federated learning provides the infrastructure through which algorithm updates agreed upon in the PCCP can be implemented across all sites without requiring patient data transfer—converting what would otherwise be a regulatory and logistical obstacle into a routine, pre-authorized model maintenance cycle.

6. Future directions and challenges

6.1. Validation standards

The most pressing unresolved challenge for AI-assisted ultrasound in clinical trials is the absence of agreed minimum standards for external validity. The evidence hierarchy described in Section 3.2 provides a conceptual framework, but trial sponsors currently have no single authoritative checklist specifying what degree of external

validation is sufficient before an AI measurement algorithm is adopted to ascertain a primary trial endpoint. Translating STARD-AI and PROBAST+AI principles into a concise, endpoint-specific validation template could lower the barrier to regulatory acceptance and reduce redundant validation work across independent trial programs (*14,15*).

6.2. Data infrastructure and foundation models

The federated foundation model paradigm exemplified by UltraFedFM (*9*) suggests a potentially important direction for ultrasound AI infrastructure. Rather than developing a separate model de novo for each trial, in the future AI-assisted ultrasound may increasingly be adapted from large pre-trained models that are subsequently fine-tuned for specific clinical or trial applications and which are then used to ascertain endpoints. If validated across endpoint-specific use cases, such an approach could reduce the amount of task-specific labelled data required and may improve cross-site robustness by capturing a broader range of scanner, operator, and population variability during pre-training. However, whether these advantages translate consistently to regulated trial endpoints still needs to be established. Realizing this potential would likely require expansion of federated training networks, with broader representation of community hospitals and LMICs. More diverse participation could help reduce the risk that foundation models reflect predominantly tertiary-center patient populations, workflows, or equipment characteristics.

The public data ecosystem also requires investment. The SonoDQS catalog identified 72 public datasets as of late 2025, but fetal and prostate ultrasound remain substantially underrepresented relative to the frequency of their clinical use (*13*). Enhancing coordinated data-sharing initiatives—and particularly those that include standardized acquisition metadata, demographic descriptors, and reference-standard labels—could facilitate both more reproducible algorithm development and more credible independent benchmarking.

6.3. Limitations of this review

Several limitations should be acknowledged. First, this is a narrative review and does not claim to represent a systematic or exhaustive survey of the literature; papers were selected to illustrate key methodological and regulatory principles, and the evidence presented may not reflect the full breadth of published findings. Second, the clinical validation evidence is weighted toward AI-assisted echocardiography, reflecting the current state of the field: rigorous multi-center prospective studies and RCTs mostly involve cardiac imaging, while evidence levels are lower for fetal, hepatic, musculoskeletal, and vascular ultrasound. Conclusions about the regulatory and methodological readiness of AI endpoints in non-cardiac domains should therefore be interpreted as anticipatory rather than established. Third, this review draws primarily on English-language literature and published regulatory guidance from the United States, European Union, and Japan; evidence and frameworks from other jurisdictions—including China, South Korea, and Brazil, each of which has issued AI medical device guidance—were not systematically covered. Fourth, the regulatory landscape is changing rapidly: specific timelines, guidance documents, and device inventories cited here reflect the state of affairs as of early 2026, and readers should verify their current status through primary regulatory sources before relying on these details for submission purposes.

7. Conclusion

AI-assisted ultrasound is advancing from research proof-of-concept toward operational use as a clinical trial tool, with a growing body of prospective evidence and formalized guidance supporting this transformation. The three evidence domains reviewed here—automated endpoint measurement, AI-guided ultrasound acquisition in decentralized settings, and regulatory and methodological readiness—each now rest on more than speculative promise, but routine adoption is limited in most non-cardiac ultrasound applications.

For investigators designing trials with ultrasound-based endpoints, the practical implications are increasingly clear. Automated measurement algorithms should be validated against a pre-specified equivalence or performance criterion before using them to ascertain primary endpoints, and subgroup performance and domain shift analyses should be reported transparently in accordance with frameworks such as STARD-AI or PROBAST+AI. AI-guided ultrasound acquisition protocols should be evaluated in the intended operator population - not merely in trained sonographers - before adoption in DCT designs. Regulatory interactions should consider change-management planning from the outset so that future algorithm modification is handled through a prospectively defined process rather than an avoidable protocol-amendment crisis.

For regulators and guideline developers, the most productive near-term investment is in harmonization: of cross-vendor benchmark standards, of endpoint validation templates, and of data sharing norms that shape public ultrasound resources to meet clinical need. These convergences will not occur spontaneously; they require coordinated action across the research, industry, and regulatory communities that this review addresses collectively.

Funding: This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (24K14216).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Shen YT, Chen L, Yue WW, Xu HX. Artificial intelligence in ultrasound. Eur J Radiol. 2021; 139:109717.
2. Thavendiranathan P, Grant AD, Negishi T, Plana JC, Popović ZB, Marwick TH. Reproducibility of echocardiographic techniques for sequential assessment of left ventricular ejection fraction and volumes: Application to patients undergoing cancer chemotherapy. J Am Coll Cardiol. 2013; 61:77-84.
3. Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, Heidenreich PA, Harrington RA, Liang DH, Ashley EA, Zou JY. Video-based AI for beat-to-beat assessment of cardiac function. Nature. 2020; 580:252-256.
4. He B, Kwan AC, Cho JH, *et al*. Blinded, randomized trial of sonographer versus AI cardiac function assessment. Nature. 2023; 616:520-524.
5. van Sloun RJG, Cohen R, Eldar YC. Deep learning in ultrasound imaging. Proc IEEE. 2020; 108:11-29.
6. Trost B, Rodrigues L, Ong C, *et al*. Artificial intelligence empowers novice users to acquire diagnostic-quality echocardiography. JACC Adv. 2025; 4:102005.
7. U.S. Food and Drug Administration. Conducting Clinical Trials With Decentralized Elements: Guidance for Industry, Investigators, and Other Interested Parties. FDA; September 2024. *https://www.fda.gov/regulatory-information/search-fda-guidance-documents/conducting-clinical-trials-decentralized-elements* (accessed April 1, 2026).
8. Vafaeezadeh M, Behnam H, Gifani P. Ultrasound image analysis with vision transformers — Review. Diagnostics. 2024; 14:542.
9. Jiang Y, Feng CM, Ren J, *et al.* From pretraining to privacy: Federated ultrasound foundation model with self-supervised learning. NPJ Digit Med. 2025; 8:714.
10. Tromp J, Bauer D, Claggett BL, *et al*. A formal validation of a deep learning-based automated workflow for the interpretation of the echocardiogram. Nat Commun. 2022; 13:6776.
11. UUSIC25 Challenge Consortium. Diagnostic performance of universal-learning ultrasound AI across multiple organs and tasks. arXiv. 2025. arXiv:2512.17279.
12. Mendez M, Jafarpisheh N, Demello S, Lee C, Dang M, Tyrrell PN. Evaluating the impact of sex bias on AI models in musculoskeletal ultrasound of joint recess distension. PLoS One. 2025; 20:e0332716.
13. Alsharid M, Guo X, Men Q, *et al*. On the public dissemination and open sourcing of ultrasound resources, datasets and deep learning models. NPJ Digit Med. 2025; 8:777.
14. Sounderajah V, Guni A, Liu X, Collins GS, Karthikesalingam A, Markar SR, Golub RM, Denniston AK, Shetty S, Moher D, Bossuyt PM, Darzi A, Ashrafian H. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. Nat Med. 2025; 31:3283-3289.
15. Moons KGM, Damen JAA, Kaul T, *et al*. PROBAST+AI: An updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ. 2025; 388:e082505.
16. Vasey B, Nagendran M, Campbell B, *et al.* Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022; 28:924-933.
17. Lekadir K, Frangi AF, Porras AR, *et al.* FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ. 2025; 388:e081554.
18. Baloescu C, McNaughton CD, Arntfield R, *et al*. Artificial intelligence-guided lung ultrasound by nonexperts. JAMA Cardiol. 2025; 10:245-253.
19. Speranza G, Mischkewitz S, Al-Noor F, Kainz B. Value of clinical review for AI-guided deep vein thrombosis diagnosis with ultrasound imaging by non-expert operators. NPJ Digit Med. 2025; 8:135.
20. U.S. Food and Drug Administration. Clinical Trial Imaging Endpoint Process Standards: Guidance for Industry. FDA; 2018. *https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-trial-imaging-endpoint-process-standards-guidance-industry* (accessed April 1, 2026).
21. Wong A, Roslan NL, McDonald R, *et al*. Clinical obstacles to machine-learning POCUS adoption and system-wide AI implementation (The COMPASS-AI survey). Ultrasound J. 2025;17:32.
22. Kim S, Fischetti C, Guy M, Hsu E, Fox J, Young SD. Artificial intelligence (AI) applications for point of care ultrasound (POCUS) in low-resource settings: A scoping review. Diagnostics. 2024; 14:1669.
23. U.S. Food and Drug Administration. Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions: Guidance for Industry and FDA Staff. FDA; December 2024. *https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-intelligence* (accessed April 1, 2026).
24. European Parliament and Council of the European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Off J Eur Union. 2024; L 2024/1689. *https://eur-lex.europa.eu/eli/reg/2024/1689/oj* (accessed April 1, 2026).
25. U.S. Food and Drug Administration, Health Canada, Medicines and Healthcare products Regulatory Agency. Good Machine Learning Practice for Medical Device Development: Guiding Principles. FDA; October 2021. *https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles* (accessed April 1, 2026).
26. Stringer JSA, Pokaprakarn T, Prieto JC, *et al*. Diagnostic accuracy of an integrated AI tool to estimate gestational age from blind ultrasound sweeps. JAMA. 2024; 332:649-657.
27. Slimani S, Hounka S, Mahmoudi A, *et al.* Fetal biometry and amniotic fluid volume assessment end-to-end automation using deep learning. Nat Commun. 2023; 14:7047.
28. Zhang L, Zhao X, Song L, *et al*. Paired liver-spleen high-frequency ultrasound deep learning network for full-stage liver fibrosis classification and clinical benefit compared with 2D-SWE in chronic hepatitis B cohort: A prospective

multicenter study. J Gastroenterol. 2026; 61:184-194.
29. Ammitzbøll-Danielsen M, Østergaard M, Tamm L, Terslev L. Diagnostic performance and reliability of robotic ultrasonography and artificial intelligence-driven synovitis assessment in rheumatoid arthritis: Results from the Controlled ARTHUR Trial. RMD Open. 2025; 11:e006099.
30. International Medical Device Regulators Forum. Good machine learning practice for medical device development: Guiding principles. IMDRF; 2025. *https://www.imdrf.org/documents/good-machine-learning-practice-medical-device-development-guiding-principles* (accessed April 1, 2026).
31. Gomes RG, Vwalika B, Lee C, *et al.* A mobile-optimized artificial intelligence system for gestational age and fetal malpresentation assessment. Commun Med. 2022; 2:128.
32. Chiu IM, Chen TY, Zheng YC, Lin XH, Cheng FJ, Ouyang D, Cheng CY. Prospective clinical evaluation of deep learning for ultrasonographic screening of abdominal aortic aneurysms. NPJ Digit Med. 2024; 7:282.
33. U.S. Food and Drug Administration. Artificial Intelligence-enabled Medical Devices. *https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-enabled-medical-devices* (accessed April 1, 2026).
34. Pharmaceuticals and Medical Devices Agency. Regulations and Approval/Certification of Medical Devices in Japan. *https://www.pmda.go.jp/english/review-services/reviews/0004.html* (accessed April 1, 2026).
35. Pharmaceuticals and Medical Devices Agency. Software as a Medical Device (SaMD). *https://www.pmda.go.jp/english/review-services/reviews/0009.html* (accessed April 1, 2026).
36. Pharmaceuticals and Medical Devices Agency. SAKIGAKE Designation System. *https://www.pmda.go.jp/english/review-services/reviews/advanced-efforts/0001.html* (accessed April 1, 2026).

Received March 10, 2026; Revised April 12, 2026; Accepted April 18, 2026.

**Address correspondence to:*
Kenji Karako, Department of Surgery, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo, Japan 113-8655.
E-mail: tri.leafs@gmail.com

Released online in J-STAGE as advance publication April 20, 2026.

Review

DOI: 10.5582/ddt.2026.01010

Classical traditional Chinese medicine formulas for inflammatory bowel disease: Therapeutic evidence and mechanistic insights

Lichao Yang[1,2,§], Zhixian Jiang[1,§], Qi Sun[1], Lianwen Yuan[1,*], Wei Tang[2]

[1] Department of General Surgery, The Second Xiangya Hospital of Central South University, Changsha, China;
[2] National Center for Global Health and Medicine, Japan Institute for Health Security, Tokyo, Japan.

SUMMARY: Inflammatory bowel disease (IBD) is a digestive system disorder characterized by chronic recurrent inflammation, primarily including ulcerative colitis (UC) and Crohn's disease (CD). Although modern precision medicine therapies, represented by biologics and small-molecule drugs, have reshaped the IBD treatment landscape, clinical practice still faces challenges such as primary non-response, secondary loss of response, risks of severe opportunistic infections, and an increasingly heavy health-economic burden. Traditional Chinese Medicine (TCM), as a unique medical system with millennia of clinical experience and guided by its distinct theoretical framework of "holistic concept" and "treatment based on syndrome differentiation," shows great potential in IBD management. In recent years, advances in evidence-based medicine and modern biotechnology have propelled TCM research in IBD from empirical treatment toward precision medicine. TCM exerts its effects through multi-target and multi-pathway mechanisms, demonstrating unique advantages particularly in immunomodulation, intestinal barrier repair, and gut microbiota regulation. Furthermore, modern research, from holistic perspectives such as the "gut-lung axis" and "gut-brain axis," scientifically interprets the modern biological connotations of traditional TCM theories, providing new understanding for TCM's application in IBD. Concurrently, this article discusses the revolutionary progress of nanodrug delivery systems and multi-omics technologies in enhancing the bioavailability of TCM components, reducing toxicity, and elucidating complex mechanisms. It aims to construct a modern scientific framework for TCM in treating IBD through the dual macroscopic and microscopic perspectives of systems biology and translational medicine, offering insights for achieving precision therapy, whole-life-cycle management of IBD, and complementary integration of Chinese and Western medicine.

Keywords: Inflammatory bowel disease, traditional Chinese medicine, immunomodulation, intestinal barrier, gut microbiota

1. Introduction

Inflammatory bowel disease (IBD) is a chronic, relapsing inflammatory disorder of the gastrointestinal tract, driven by complex interactions between genetic susceptibility, immune dysregulation, environmental factors, and disturbances in gut microbiota that disrupt intestinal homeostasis (*1*). IBD primarily includes ulcerative colitis (UC) and Crohn's disease (CD). UC mainly involves the colon and rectum, presenting as continuous mucosal inflammation (*2*); whereas CD can affect the entire gastrointestinal tract, exhibiting skip lesions and transmural inflammation (*3*). IBD was long considered a disease specific to Western industrialized nations. However, accelerated globalization, profound dietary Westernization, antibiotic overuse, and environmental changes have led to its rapid spread in emerging economies across Asia, South America, and Africa, evolving it into a global epidemic (*4-7*) . This profound epidemiological shift further underscores the central role of complex gene-environment interactions in disease pathogenesis (*8*). According to the Global Burden of Disease (GBD) 2021 estimates, there are approximately 3.83 million cases of chronic inflammatory bowel disease worldwide, with 375,000 new cases and 42,000 deaths in 2021, resulting in more than 1.5 million disability-adjusted life years (DALYs) lost (*9*). IBD significantly impacts patients' quality of life, with long-term and severe symptoms contributing to substantial social, emotional, and economic burdens. Studies show that the prevalence of anxiety and depression among IBD patients is 35.1% and 21.6%, respectively (*10*). Furthermore, the long-term complications of IBD, including colorectal cancer (CRC) (with a 7% risk for patients with 30 years of disease duration) (*11*), intestinal strictures, fistula formation, and extra-intestinal manifestations (such as

arthritis and liver disease), increase the risk of disability and mortality. Among Crohn's disease (CD) patients, approximately 33% will develop intestinal strictures, with 7.8% having multiple strictures; 10.6% will develop fistulas, with perianal fistulas being the most common. Long-term follow-up studies indicate that around 50% of CD patients will experience strictures and/or fistulas within 20 years of diagnosis, significantly impacting quality of life and increasing clinical resource consumption (*12*).

Current therapeutic goals have evolved from mere clinical symptom remission to endoscopic mucosal healing and even deeper histological healing. Although modern medicine has achieved milestone progress in IBD treatment — from conventional 5-aminosalicylic acid (5-ASA), corticosteroids, and immunomodulators to biologics and small-molecule drugs — limitations remain in clinical practice. Some patients exhibit primary non-response to biologics or develop secondary loss of response due to anti-drug antibodies. Furthermore, the risks of infection and malignancy associated with long-term immunosuppressive therapy cannot be ignored (*13*). Although these treatments improve clinical outcomes, they do not provide a cure, and their side effects limit their long-term use. Additionally, the high cost of biologic treatments and their accessibility remain a critical issue, particularly in low-resource settings (*14*). Therefore, novel, safer, and more effective therapies are urgently needed to address the remaining gaps in the management of IBD.

Traditional Chinese Medicine (TCM), a treasure of Chinese civilization, has a rich theoretical foundation in the treatment of IBD. UC is often categorized as "chronic dysentery". Its pathogenesis is generally characterized by the intermingling of deficiency and excess, with damp-heat accumulation in the large intestine and qi stagnation regarded as major pathogenic factors (*15*). The pathogenesis of CD is more complex, often involving intermingled cold and heat, deficiency and excess, and is prone to forming intestinal strictures ("accumulation") as the disease progresses (*15,16*). Unlike modern medicine's focus on "antagonistic" therapy blocking single or specific inflammatory pathways, TCM emphasizes "harmonization," restoring the body's immune homeostasis, microecological balance, and barrier integrity through the synergistic action of multiple components, targets, and levels. Recently, with the rigorous introduction of evidence-based medicine methodologies and rapid development of modern biotechnology, TCM research in IBD has successfully transcended the boundaries of empirical medicine, entering a new stage of mechanistic elucidation and precision translation. This is not only due to numerous high-quality randomized controlled trials (RCTs) confirming the efficacy of Chinese herbs in inducing and maintaining remission but also because basic research has deeply revealed their fine regulatory mechanisms at the molecular and cellular levels (*17,18*). Furthermore, research and application of nanodrug delivery systems can significantly enhance the solubility and bioavailability of TCM components, reduce side effects through targeted delivery, opening new prospects for the modernization and internationalization of TCM.

This article will review and summarize the application and research progress of TCM formulas in IBD treatment, focusing on their mechanisms of action in regulating the immune microenvironment, intestinal barrier function, and gut microbiota balance. It aims to provide theoretical support for the evidence-based application of TCM in IBD treatment and offer a scientific basis for future precision therapy. A comprehensive literature search was conducted across multiple databases, including PubMed, Web of Science, MEDLINE, EMBASE, Springer LINK, Wanfang Database, China Biomedical Literature Database, and China National Knowledge Infrastructure (CNKI). The keywords used for the search were "inflammatory bowel disease" (including "ulcerative colitis" or "Crohn's disease") and "Traditional Chinese Medicine". All selected articles were published between 2016 and 2026.

2. New evidence for clinical efficacy and application strategies of classical TCM formulas

TCM formulas are the core vehicle of TCM treatment, designed according to the rigorous compatibility principle of "sovereign, minister, assistant, and envoy" ("Jun Chen Zuo Shi"), embodying the essence of systematic TCM therapy. TCM formulas holds unique advantages in treating IBD, manifested through synergistic regulation of multiple targets and pathways (Figure 1). Its application strategies are flexible and diverse, serving as independent therapy for mild-to-moderate patients or as adjuvant therapy for severe or refractory cases to enhance Western drug efficacy and reduce side effects (Table 1). Its clinical value has been validated through long-term practice, and recent evidence from evidence-based medicine has further solidified its therapeutic efficacy.

2.1. Shenling Baizhu San

Shenling Baizhu San, originating from the Taiping Huimin Heji Ju Fang in the Song Dynasty, consists of ginseng, Atractylodes, Poria, Chinese yam, lotus seed, hyacinth bean, coix seed, Amomum, Platycodon, and licorice. It is a classic formula for treating "Spleen Deficiency and Dampness Stagnation," commonly used in the clinical practice of IBD, particularly during the remission phase or for mild to moderate active IBD. It is effective in improving symptoms such as diarrhea, fatigue, and poor appetite (*19,20*). In the context of IBD, especially UC, this formula is explicitly recommended as the first choice for treating the spleen deficiency

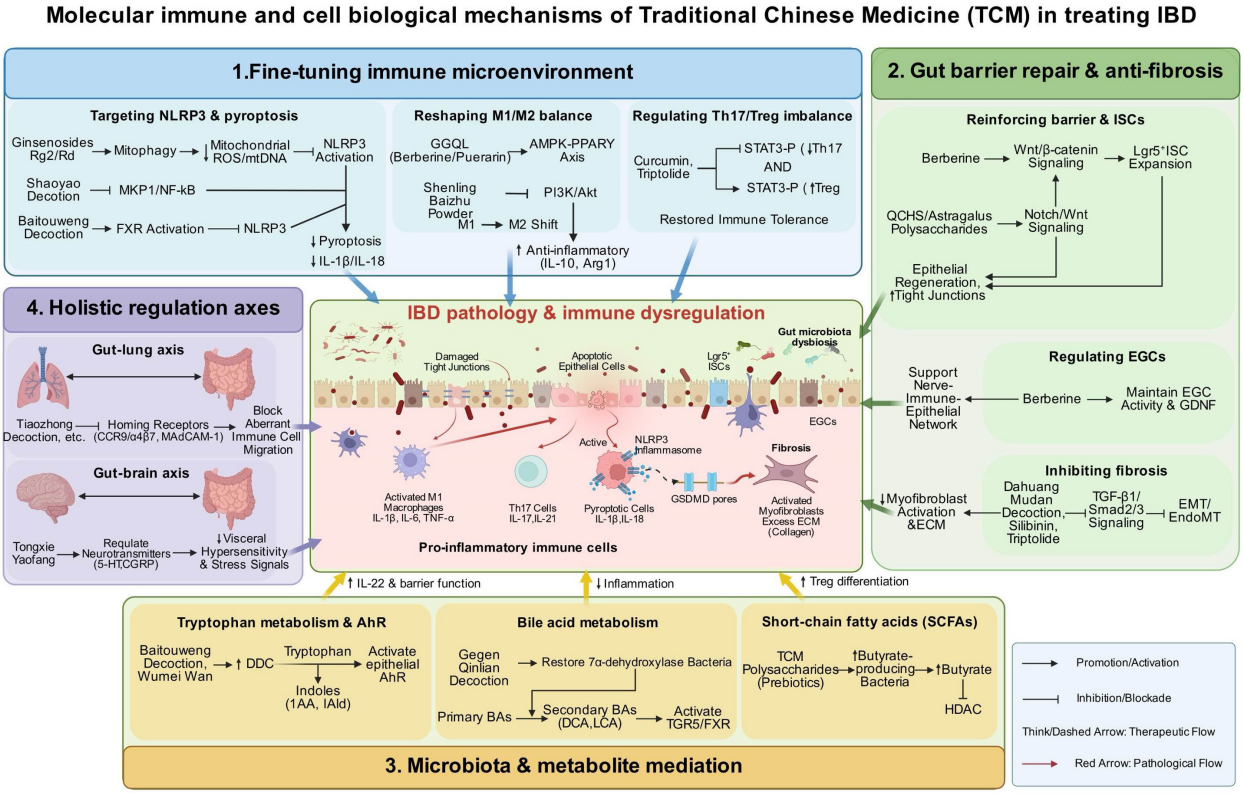

Figure 1. Molecular immune and cell biological mechanisms of traditional Chinese medicine (TCM) in treating inflammatory bowel disease (IBD). This figure depicts the multifaceted therapeutic actions of TCM in IBD, organized into four core modules: 1. Immune microenvironment fine-tuning: TCM modulates NLRP3 inflammasome/pyroptosis, M1/M2 macrophage balance, and Th17/Treg tolerance to suppress inflammation. 2. Gut barrier repair & anti-fibrosis: TCM promotes Lgr5+ ISC expansion and epithelial regeneration, sustains EGC activity, and inhibits myofibroblast activation/ECM deposition to mitigate fibrosis. 3. Microbiota & metabolite mediation: TCM regulates gut microbiota as well as the metabolism of tryptophan, bile acids, and short-chain fatty acids to enhance barrier function and reduce inflammation. 4. Holistic regulation: TCM suppresses aberrant immune cell trafficking by regulating homing receptors (CCR9/CCL25 and α4β7/MAdCAM-1) in the gut-lung axis, and modulates neurotransmitters (5-HT, CGRP) to attenuate visceral hypersensitivity in the gut-brain axis, exerting holistic therapeutic effects.

and dampness stagnation type according to diagnostic and therapeutic guidelines (*21,22*), and is suitable for mild to moderate active phase and remission phase UC, especially for managing chronic relapsing UC.

Recent evidence from evidence-based medicine has further confirmed the clinical efficacy and synergistic value of Shenling Baizhu San (*23*). For example, a meta-analysis published in China Pharmacy (including 14 RCTs with 1,177 UC patients) demonstrated that the combination of Shenling Baizhu San with 5-aminosalicylic acid (5-ASA) significantly improved the total response rate compared to 5-ASA alone, reducing the disease activity index (DAI) and lowering pro-inflammatory cytokines such as TNF-α and IL-17 (*24*). The 2023 Guidelines for the Integrated Diagnosis and Treatment of UC further consolidated the evidence, confirming the advantages of combining Shenling Baizhu San with Western medications in improving clinical efficacy, reducing DAI scores, and alleviating symptoms such as abdominal pain and diarrhea. Additionally, recent randomized controlled trials (RCTs) have further verified its efficacy. For instance, a domestic study involving 100 UC patients, a study including 62 pediatric and adolescent UC patients, and another on 70 elderly UC patients with bloody diarrhea all confirmed the positive role of Shenling Baizhu San in UC treatment (*25-27*). A more comprehensive prospective cohort study (48 patients) showed that 8 weeks of Shenling Baizhu San combined therapy not only significantly improved clinical symptoms and endoscopic mucosal healing but also verified its effects through 16S sequencing and metabolomics, showing close associations with gut microbiota remodeling and increased levels of tryptophan metabolites (such as IPA, IAA) (*28*). This finding clarifies the key mechanism of the "microbiota-tryptophan metabolic axis" in the synergistic effect of Shenling Baizhu San. Overall, Shenling Baizhu San demonstrates multiple clinical advantages in UC treatment, including symptom relief, disease activity reduction, mucosal healing, and providing a new treatment mechanism through gut microbiota and metabolic pathway regulation.

2.2. Shaoyao Tang

Shaoyao Tang, originating from the "Su Wen Bing

Table 1. Representative traditional Chinese medicine formulas for inflammatory bowel disease: corresponding TCM syndromes, core molecular mechanisms, and key clinical/experimental effects

Classical TCM formulas	Corresponding TCM syndrome	Major active components	Core molecular mechanisms	Key clinical/experimental effects	Ref
Shenling Baizhu San	Spleen deficiency with dampness retention	Ginsenosides, atractylenolides, Poria polysaccharides	Activation of the AhR–CYP1A1–NF-κB pathway; regulation of the SLC6A14/PI3K–AKT axis; improvement of the comorbidity network linking insulin resistance and intestinal inflammation	Enhancement of intestinal barrier integrity; significant reduction of inflammatory and metabolic indicators in patients with T2DM complicated by UC; regulation of Th17/Treg balance	(*19,20,28,67*)
Shaoyao Decoction	Damp-heat in the large intestine	Paeoniflorin, berberine	Blockade of the MKP1/NF-κB pathway; inhibition of NLRP3 inflammasome assembly and activation; suppression of ferroptosis in intestinal epithelial cells; improvement of hypercoagulable status	Significant reduction of serum TNF-α, IL-1β, and IL-6 levels; improvement of microcirculation; alleviation of DSS-induced colonic injury and bleeding	(*31,44*)
Baitouweng Decoction	"Heat-Toxin Excess" and "Damp-Heat Accumulation"	Pulsatilla saponins, berberine, fraxin	Upregulation of dopa decarboxylase (DDC), promoting tryptophan metabolism toward the indole pathway; increased IAA production and activation of the AhR/IL-22 axis; regulation of bile acid metabolism and activation of FXR	Promotion of deep mucosal healing; restoration of gut microbiota diversity; reduction of DAI scores; efficacy in refractory UC.	(*34-37,45,59*)
Tongxie Yaofang	Liver Qi stagnation and Spleen deficiency	Cimifugin glycoside, 5-O-methylvisamminol glycoside	Regulation of the brain–gut axis; reduction of visceral hypersensitivity; modulation of 5-HT, CGRP, and BDNF levels; inhibition of the RAGE signaling pathway	Alleviation of abdominal pain, diarrhea, and psychological symptoms; improvement of quality of life in patients with IBS-D and mild UC; regulation of stress-induced colitis	(*39,69,70*)
Gegen Qinlian Decoction	Damp-heat diarrhea	Puerarin, berberine, baicalin	Remodeling of the microbiota–bile acid axis; enrichment of 7α-dehydroxylase–producing bacteria; increased secondary bile acids (DCA, LCA); activation of TGR5/VDR receptors	Restoration of goblet cell function; promotion of mucin MUC2 secretion; improvement of intestinal inflammation associated with dysregulated glucose and lipid metabolism	(*40-42,61,82*)

Ji Qi Yi Bao Ming Ji", is based on the core functions of "clearing heat and drying dampness, regulating Qi and promoting blood circulation." It consists of nine herbs: white peony root, Chinese angelica root, Coptis root, areca nut, costus root, rhubarb root, Baikal skullcap root, cinnamon bark, and licorice root. This classic formula is used in the treatment of "Damp-Heat Accumulation in the Large Intestine" and has been strongly recommended in the 2023 Guidelines for the Integrated Diagnosis and Treatment of UC for mild to moderate active UC, especially those presenting with abdominal pain, tenesmus, and bloody diarrhea. It is also strongly recommended (Grade III evidence) as an adjunctive therapy in moderate to severe UC when used in combination with biologics (*21,29,30*).

Multiple evidence-based studies have confirmed its synergistic value. The latest meta-analysis (including 23 RCTs) demonstrated that Shaoyao Tang, whether used alone or in combination with Western medicines, significantly improved clinical efficacy, symptom efficacy, and quality of life (IBDQ scores), while also reducing recurrence rates and decreasing levels of pro-inflammatory cytokines such as TNF-α, IL-1β, and IL-6, and increasing the expression of anti-inflammatory cytokines such as IL-4 and IL-10 (*31*). The 2023 guidelines further summarize this evidence, confirming the combination of Shaoyao Tang with mesalazine in mild to moderate UC improves clinical remission rates, endoscopic response rates, and mucosal healing rates (*30*). RCTs of different disease stages further verify its efficacy: in mild to moderate UC (127 patients, RCT), Shaoyao Tang combined with mesalazine for 4 weeks showed a clinical remission rate of approximately 80%, significantly better than mesalazine alone, and significantly improved symptom scores for abdominal pain and diarrhea, as well as Baron endoscopic scores (*32*); in moderate to severe UC (68 patients, RCT), infliximab combined with Shaoyao Tang for 12 weeks showed superior clinical efficacy, Mayo scores, and endoscopic scores compared to infliximab alone (*33*). Moreover, Shaoyao Tang is considered to have good overall safety, with most RCTs reporting no significant adverse reactions. Meta-analysis results show that the incidence of adverse events in the combination therapy group was not higher than that in the monotherapy group, and some studies even suggested that combined therapy could significantly reduce adverse events, with no severe adverse events reported, though long-term safety still needs further verification.

2.3. Baitouweng Decoction

Baitouweng Decoction, originating from the Shang Han Lun, is based on the core functions of "clearing heat and detoxifying, cooling the blood, and stopping dysentery." It consists of four herbs: Pulsatilla root, Coptis root, Phellodendron bark, and Ash bark. This formula is a classic treatment for "Heat-Toxin Excess" and "Damp-Heat Accumulation" type UC. It is explicitly listed as a core formula for the active phase of UC in both the *Expert Consensus on the Diagnosis and Treatment of UC in Traditional Chinese Medicine (2017)* and the *Expert Consensus on the Diagnosis and Treatment of UC in Traditional Chinese Medicine (2023)*, particularly for patients with prominent symptoms of abdominal pain, diarrhea, and bloody stools, and is commonly administered orally or as a retention enema in combination with Western medications (*21*).

Multiple evidence-based studies have confirmed the significant efficacy of Baitouweng Decoction in UC treatment. A 2024 meta-analysis (including 24 RCTs with 2,131 patients) showed that Baitouweng Decoction, whether used alone or in combination with conventional Western medications, significantly improved clinical efficacy, reduced TCM syndrome scores and Baron endoscopic scores, and improved the inflammatory factor profile, with significant decreases in serum TNF-α and IL-8 levels, and an increase in anti-inflammatory IL-10 levels (*34*). Further clinical studies have verified its synergistic effect. A prospective cohort study (50 patients) demonstrated that Baitouweng Decoction combined with Western medicines for 30 days significantly improved clinical remission rates (69.2% *vs.* 37.5%) and total effective rates (84.6% *vs.* 58.3%) compared to Western medicines alone, with better reductions in ESR and fecal calprotectin levels and no significant adverse reactions (*35*); in another RCT (44 patients), Baitouweng Decoction combined with mesalazine for 30 days resulted in greater improvements in clinical efficacy and IBDQ quality of life scores compared to mesalazine alone, with significant increases in $CD4^+/CD8^+$ ratio and NK cell levels, and more significant reductions in TNF-α, IL-17, and IL-23 (*36*). A large sample RCT further confirmed that Baitouweng Decoction combined with mesalazine was superior in intestinal mucosal barrier repair, CRP, and mucin improvements, with a lower relapse rate compared to mesalazine alone (*37*). The overall safety of this formula is also good, with very few adverse events, mostly mild gastrointestinal discomfort such as nausea, and no severe adverse reactions reported. Meta-analysis results indicate no significant difference in adverse event rates between the combined therapy and monotherapy groups.

2.4. Tongxie Yaofang

Tongxie Yaofang, originating from Danxi Xinfa, consists of fried Atractylodes rhizome, fried white peony root, tangerine peel, and Saposhnikovia root. Its core functions include "soothing the liver, strengthening the spleen, dispelling dampness, and stopping diarrhea," making it a classic formula for treating "Liver Qi stagnation and Spleen deficiency" type ulcerative colitis (UC) in TCM. This formula is clearly recommended in relevant

guidelines and expert consensus, particularly for patients with mild to moderate active UC, especially those presenting with abdominal pain, diarrhea, and stress-related symptoms. It is commonly used in combination with mesalazine or other Western medications to enhance therapeutic efficacy (*21,30*).

In recent years, clinical studies based on randomized controlled trials (RCTs) have confirmed the clinical efficacy of Tongxie Yaofang. For example, a multicenter RCT including 40 UC patients with liver-spleen disharmony type in the active phase showed that after 12 weeks of combined treatment with mesalazine and Tongxie Yaofang, the experimental group significantly outperformed the control group (treated with mesalazine alone) in terms of the Sutherland index, Baron mucosal scores, and modified Mayo scores. Moreover, immune function markers (such as CD4+, CD4/CD8 ratio, and NK cell levels) and intestinal mucosal barrier function (with upregulation of tight junction proteins and downregulation of β-defensin) were significantly improved in the experimental group, demonstrating that Tongxie Yaofang enhances therapeutic outcomes through immune and mucosal barrier regulation (*38*). Additionally, several similar RCTs have reported that the combination therapy group showed significantly greater reductions in core symptoms such as abdominal pain, diarrhea, and bloody stools compared to the mesalazine-only group. This formula is also a representative remedy for regulating the "gut-brain axis." Research has confirmed its ability to reduce visceral hypersensitivity, regulate both central and peripheral serotonin (5-HT) levels, and modulate gut immune-epithelial communication through the IL-10RA/NF-κB pathway, alleviating stress-induced intestinal inflammation (*39*).

2.5. Gegen Qinlian Tang

Gegen Qinlian Tang is a classic TCM formula composed of Pueraria root, Baikal skullcap root, Coptis root, and roasted licorice root. Its core functions are "clearing heat, draining dampness, stopping diarrhea, and treating dysentery," and it is widely used to treat damp-heat type ulcerative colitis (UC). In clinical practice, Gegen Qinlian Tang is often used as an adjunctive therapy for UC patients with damp-heat syndrome, administered either orally or as a retention enema.

The clinical efficacy of Gegen Qinlian Tang has also been supported by relevant RCTs and meta-analysis evidence. An RCT including 60 patients showed that after 8–12 weeks of combined treatment with Gegen Qinlian Tang and mesalazine, the total effective rate was significantly higher than that of mesalazine alone (96.66% *vs.* 60.00%), with a reduction in the incidence of adverse effects. Additionally, combination therapy significantly lowered inflammatory markers such as IL-8, TNF-α, and CRP (*40*). Another RCT involving 120 patients with damp-heat type UC employed a modified version of Gegen Qinlian Tang combined with acupuncture and mesalazine/pefloxacin for 1 month. The total effective rate also increased significantly, and the combination group showed superior improvements in the Geboes index, Baron score, and IBDQ score, with a lower relapse rate at 6 months (*41*). The latest meta-analysis further confirmed that combination therapy improves the overall effective rate (RR = 1.22), reduces the adverse event rate (RR = 0.59), and significantly improves the modified Mayo score and Baron endoscopic score (*42*). The adverse effects of Gegen Qinlian Tang are mainly mild gastrointestinal discomfort, with no reports of serious adverse reactions, demonstrating good tolerance in treating damp-heat type UC. Therefore, it is commonly recommended by guidelines as an adjunctive therapy for damp-heat type UC (*30*).

3. Molecular immunology and cellular biology mechanisms of TCM in treating IBD

3.1. Fine-tuned regulation of the immune microenvironment

The immune microenvironment plays a pivotal role in the initiation and perpetuation of IBD. Immune dysregulation in IBD is mainly manifested by aberrant activation of immune cells, imbalanced cytokine secretion, and disruption of intestinal barrier integrity. Increasing evidence suggests that modulation of the immune microenvironment — particularly through key mechanisms such as NLRP3 inflammasome activation, macrophage polarization, and Th17/Treg imbalance — can provide novel theoretical support and therapeutic strategies for IBD management.

3.1.1. Targeting the NLRP3 inflammasome and pyroptosis

The NLRP3 inflammasome is a core sensor of the innate immune system. Its excessive activation leads to caspase-1–mediated cleavage of gasdermin D, thereby inducing pyroptosis and massive release of pro-inflammatory cytokines such as IL-1β and IL-18. The excessive production of these cytokines is considered one of the major driving forces of the inflammatory storm in IBD. Multiple active components derived from TCM have been shown to specifically inhibit NLRP3 inflammasome activation. For instance, ginsenosides Rg2 and Rd promote mitophagy, facilitating the clearance of damaged mitochondria and reducing the release of mitochondrial reactive oxygen species (ROS) and mitochondrial DNA (mtDNA), thereby blocking NLRP3 activation at its source (*43*). Shaoyao Decoction suppresses the upstream MKP1/NF-κB signaling pathway, resulting in reduced transcription of NLRP3 and pro-IL-1β (*44*). In contrast, Baitouweng Decoction modulates bile acid metabolism, activates the farnesoid X receptor (FXR), and interacts with NF-κB signaling to inhibit NLRP3 expression. This form of "metabolism–

immunity" crosstalk represents a novel perspective on TCM-based intervention, highlighting the potential of herbal medicine in regulating pyroptosis and immune-driven inflammation (*45*).

3.1.2. Remodeling macrophage polarization (M1/M2 balance)

Macrophages play a critical role in immune tolerance, and their plasticity determines the balance of the immune microenvironment. In IBD, macrophages are polarized toward the pro-inflammatory M1 phenotype, which secretes large amounts of pro-inflammatory factors and promotes disease progression. The therapeutic goal is to reprogram M1 macrophages into the anti-inflammatory and tissue-repairing M2 phenotype, thereby alleviating inflammation and promoting intestinal repair. The AMPK–PPARγ axis is a key pathway regulating M1/M2 macrophage polarization. The Gegen Qinlian (GGQL) nanoformulation contains components such as berberine and puerarin, which upregulate PPARγ expression through activation of AMPK phosphorylation. As a nuclear receptor, PPARγ can directly suppress the transcription of inflammatory genes and promote the expression of M2 markers (such as CD206, Arg1, and IL-10), thereby effectively enhancing immune tolerance and intestinal repair (*46*). Shenling Baizhu San has also been shown to reduce M1 polarization and increase the proportion of M2 macrophages by inhibiting the PI3K/Akt pathway, further promoting the repair of the intestinal mucosa. These TCM components exhibit significant effects in remodeling macrophage polarization, providing new insights into immunotherapy for IBD (*47*).

3.1.3. Regulation of Th17/Treg imbalance

Th17/Treg imbalance is a key event in the immunopathological process of IBD. Th17 cells drive intestinal mucosal inflammation by secreting pro-inflammatory cytokines such as IL-17 and IL-21, whereas Treg cells suppress excessive immune responses and maintain tolerance through IL-10 and TGF-β (*48*). The inflammatory microenvironment, STAT3/STAT5 signaling pathways, intestinal microbiota, and metabolic products jointly regulate the differentiation and function of Th17 and Treg cells. In IBD, the relative overexpression of pro-inflammatory Th17 cells and the reduction or functional impairment of suppressive Treg cells disrupt immune homeostasis, representing an important mechanism underlying persistent inflammation (*49*). Several TCM components, such as curcumin, triptolide, and Qingjie Fuzheng granules, have shown significant effects in restoring the Th17/Treg balance. These components inhibit Th17 differentiation by suppressing STAT3 phosphorylation while promoting Treg differentiation by enhancing STAT5 phosphorylation, thereby restoring immune tolerance. In this manner, curcumin and related compounds not only alleviate Th17-mediated pro-inflammatory responses but also enhance the immunosuppressive function of Treg cells, providing new immunomodulatory strategies for IBD treatment (*50*).

3.2. Repair and regeneration of the intestinal barrier

The homeostasis of the intestinal barrier is fundamental for maintaining the separation between the intestinal lumen and the internal environment of the host. Damage to the intestinal barrier is one of the key events in the pathogenesis of IBD. Repair of the intestinal barrier involves multiple aspects, including reinforcement of the mechanical barrier, regeneration of epithelial cells, neuroimmune interaction, and prevention of pathological fibrosis (*51*). Therefore, restoration of intestinal homeostasis and termination of chronic inflammation require a comprehensive repair of the intestinal barrier.

3.2.1. Reinforcement of the mechanical barrier and intestinal stem cells (ISCs)

The mechanical barrier is mainly composed of intestinal epithelial cells and their tight junction complexes (such as ZO-1 and occludin), which are usually severely disrupted during the active phase of IBD. Various TCM components can promote the expression of tight junction proteins and strengthen barrier function, thereby reducing inflammation caused by the translocation of pathogens and endotoxins (*51*). Intestinal stem cells (ISCs), located at the base of the crypts and characterized by the expression of the marker Lgr5, represent the fundamental driving force for epithelial renewal and repair. Recent studies have shown that certain Chinese patent medicines, such as berberine, not only exhibit anti-inflammatory effects but also regulate the microecological and signaling environments of ISCs, thereby promoting epithelial regeneration. Luo *et al.* found in a DSS-induced colitis model that berberine upregulated the expression of Wnt genes in resident stromal cells, enhanced the activity of the Wnt/β-catenin pathway, and promoted the activation and expansion of $Lgr5^+$ ISCs through this mechanism, thereby accelerating the repair of damaged intestinal mucosa. Blocking Wnt secretion significantly attenuated the reparative effects of berberine, indicating that Wnt/β-catenin signaling plays a critical role in this process (*52*). In addition, TCM formulas such as Qingchang Huashi Formula (QCHS) and Astragalus polysaccharides have been reported to promote the proliferation and differentiation of $Lgr5^+$ stem cells through the Notch and Wnt signaling pathways, enhancing the replenishment of multiple epithelial cell types, including goblet cells, and thereby accelerating the recovery of damaged regions (*85*). These mechanisms are consistent with findings from intestinal

organoid studies, which demonstrate the interactive roles of Wnt–Notch networks in cell fate determination, indicating that regulation of the stem cell niche is a key strategy for barrier repair (*53*).

3.2.2. Regulation of enteric glial cells (EGCs)

Enteric glial cells (EGCs), which are analogous to astrocytes in the central nervous system, constitute an important component of the intestinal microenvironment and maintain complex interaction networks with epithelial cells and the immune system. EGCs secrete regulatory factors such as neurotrophic factors (*e.g.*, GDNF), which promote epithelial cell maturation, enhance tight junction function, and strengthen antimicrobial defense, thereby supporting the integrity of the intestinal barrier (*54*). Studies have shown that in experimental colitis models, reduced expression or functional abnormalities of EGCs exacerbate barrier disruption; conversely, preservation or restoration of EGC function can alleviate inflammation and maintain barrier structure. In this context, the regulatory effects of alkaloids such as berberine on EGCs have gradually been revealed. Existing studies indicate that berberine can maintain the residence and activity of EGCs, regulate the interactions among EGCs, intestinal epithelial cells, and immune cells, and suppress the excessive release of pro-inflammatory gliotransmitters, thereby reducing epithelial injury and promoting barrier repair. The specific mechanisms may involve the modulation of inflammatory cytokines and neurotrophic factors, thereby stabilizing the immune environment and barrier function. These findings reflect the potential role of TCM in regulating the "neuro–immune–epithelial" network (*55*).

3.2.3. Inhibition of intestinal fibrosis

Intestinal fibrosis is a severe complication of IBD (especially Crohn's disease), characterized by excessive deposition of extracellular matrix (ECM) in the mucosa and intestinal wall, activation of myofibroblasts, and destruction of crypt architecture, leading to intestinal strictures and obstruction. This condition severely impairs patients' quality of life, and currently, no specific therapies are available. The TGF-β/Smad signaling pathway is the core mechanism driving fibrosis. TGF-β1 activates Smad2/3, which then forms a complex with Smad4 and translocates into the nucleus, promoting the differentiation of fibroblasts into myofibroblasts and enhancing the expression of ECM components, thereby facilitating fibrosis. Inhibition of this pathway is considered a key anti-fibrotic strategy (*56,57*).

With regard to TCM interventions, multiple formulas and monomeric compounds have demonstrated anti-fibrotic potential. Some compound formulas, such as Dahuang Mudan Decoction and Qingchang Tongluo Decoction, as well as monomeric compounds including silibinin and triptolide, have been confirmed *in vitro* and in animal experiments to significantly reduce TGF-β1–induced phosphorylation of Smad2/3, block the EMT/EndoMT process, and decrease the expression of fibrosis-related proteins (such as α-SMA and collagen), thereby inhibiting the transformation of intestinal fibroblasts into myofibroblasts and reducing ECM deposition (*58*). In addition, natural compounds (such as polyphenols and flavonoids) also exhibit anti-fibrotic effects by regulating the TGF-β/Smad, NF-κB, and ECM remodeling pathways, providing new directions for the treatment of intestinal fibrosis.

3.3. The key mediating role of gut microbiota and metabolites

The therapeutic effects of TCM *in vivo* are often exerted through its interactions with the gut microbiota. This "drug–microbiota–host" interplay constitutes an essential component of TCM efficacy. The gut microbiota not only influences drug absorption and metabolism but also regulates immune responses, intestinal barrier function, and systemic health through its metabolites, such as short-chain fatty acids, bile acids, and tryptophan-derived metabolites. Therefore, the roles of gut microbiota and metabolites in TCM-based treatment of IBD should not be overlooked.

3.3.1. Tryptophan metabolism and the aryl hydrocarbon receptor (AhR)

Tryptophan is an essential amino acid that can be metabolized by gut microbiota (such as *Lactobacillus* and *Clostridium* species) into indole derivatives, including indole-3-acetic acid (IAA) and indole-3-aldehyde (IAld). These metabolites serve as endogenous ligands for the aryl hydrocarbon receptor (AhR). AhR is a nuclear receptor that plays an important role in immune responses, mainly by regulating IL-22 secretion and intestinal barrier function, thereby contributing to the maintenance of intestinal immune homeostasis and barrier integrity. Activation of AhR promotes the repair of intestinal epithelial cells, enhances intestinal defense, and consequently suppresses intestinal inflammation. Studies have shown that TCM formulas such as Baitouweng Decoction and Wumei Pill can significantly upregulate the expression of dopa decarboxylase (DDC) in the intestine, thereby promoting the metabolism of tryptophan toward the indole pathway and increasing the production of AhR ligands. This mechanism contributes to intestinal barrier repair and alleviation of intestinal inflammatory responses (*59,60*). Through this pathway, TCM can modulate AhR signaling to exert anti-inflammatory and reparative effects, thereby facilitating the restoration of intestinal function.

3.3.2. Bile acid metabolism

Abnormal bile acid metabolism in the intestine is closely associated with the occurrence and progression of IBD. Patients with IBD often exhibit reduced levels of secondary bile acids, such as deoxycholic acid (DCA) and lithocholic acid (LCA), which leads to immune dysregulation and impairment of the intestinal barrier. Secondary bile acids exert anti-inflammatory effects by activating the TGR5 and FXR receptors, thereby promoting immune tolerance and barrier repair. Studies have found that Gegen Qinlian Decoction and its modified formulas can effectively reshape the gut microbiota structure and restore the abundance of bacteria with 7α-dehydroxylase activity, thereby promoting the conversion of primary bile acids into secondary bile acids. Through modulation of the FXR/NLRP3 signaling pathway, these formulas can suppress intestinal inflammatory responses (*61*). These findings indicate that TCM can exert anti-inflammatory effects and alleviate pathological responses in IBD patients by promoting bile acid metabolism and reshaping gut microbiota composition.

3.3.3. Short-chain fatty acids (SCFAs)

Short-chain fatty acids (SCFAs) are important metabolites produced by gut microbiota through the fermentation of dietary fiber, mainly including acetate, propionate, and butyrate. Among them, butyrate serves as a major energy source for intestinal epithelial cells and acts as a histone deacetylase (HDAC) inhibitor, regulating immune cell function at the epigenetic level, particularly in Treg cell differentiation. Polysaccharide components in TCM (such as Astragalus polysaccharides and Dendrobium polysaccharides) have been shown to possess prebiotic effects, promoting the growth of butyrate-producing bacteria (such as Faecalibacterium and Enterococcus) and increasing butyrate production. These butyrates not only provide energy for intestinal epithelial cells but also promote Treg cell differentiation through epigenetic regulation, thereby enhancing intestinal immune tolerance and modulating immune balance (*62*). Therefore, polysaccharide components in TCM exert immunomodulatory and intestinal-reparative effects by promoting the production of SCFAs.

3.4. New perspectives on holistic regulation: the "lung–gut axis" and the "brain–gut axis"

The core advantage of TCM in the treatment of IBD lies in its multidimensional regulation guided by the "holistic view." The gut-lung axis and gut-brain axis, as key inter-organ and inter-system communication networks, are essential mechanisms through which TCM achieves "local anti-inflammation and systemic homeostasis" (*63,64*). Both axes regulate the local inflammation in the gut in close association with systemic immunity, neurofunction, and gut microbiota balance, through pathways involving immune cell migration, neurotransmitter transmission, and microbiota-metabolite mediation. The precise modulation of these axes by TCM further highlights its "treating both the root and the branch" therapeutic characteristics.

3.4.1. The modern immunological basis of the theory that "the lung and the large intestine are interior–exteriorly related"

In TCM theory, the lung and the large intestine are considered to have an "interior–exterior" relationship. This concept has been reinterpreted in modern immunology, giving rise to the concept of the "gut–lung axis." This concept indicates that the lung and the intestine interact through the immune system, nervous system, and microbiota, thereby maintaining systemic immune homeostasis. The mucosal immune systems of the gut and the lung are interconnected through lymphocyte trafficking. Immune cells in the gut, under the guidance of specific chemokines, express receptors such as CCR9 and α4β7/MAdCAM-1, which bind to their corresponding ligands and thereby "home" to intestinal tissues. Similarly, immune cells in the lung also "home" to pulmonary tissues through these receptor-mediated mechanisms. Under pathological conditions, aberrant homing of immune cells may occur, leading to the migration of inflammatory cells from the lung to the intestine, thereby triggering or exacerbating enteritis, and vice versa (*65,66*). This phenomenon has been observed not only in IBD but also in other immune-related diseases.

TCM plays an important role in regulating the "gut–lung axis." Studies have shown that herbs such as Tiaozhong Decoction, Chebula (*Terminalia chebula*), and *Astragalus* polysaccharides can not only alleviate colitis but also reduce concomitant pulmonary inflammation. The underlying mechanisms may involve regulation of the expression of homing receptors such as CCR9/CCL25 and α4β7/MAdCAM-1, blocking cross-organ immune cell migration, and maintaining the balance of the microbiota in both the lung and the intestine (*68*). Related studies have demonstrated that these formulas exhibit significant immunomodulatory effects in both clinical and animal experiments, providing cellular and molecular evidence for the theory of "treating different diseases with the same method." This discovery offers new support for the use of TCM in treating diseases associated with lung–gut interactions.

3.4.2. The brain–gut axis and emotion-related pathogenesis

The brain–gut axis refers to the bidirectional communication network between the gut and the brain, involving multiple mechanisms, including neural, endocrine, and immune pathways. Recent studies have shown that patients with IBD often experience

psychological symptoms such as anxiety and depression, and psychological stress can in turn trigger disease activity. Psychological factors directly affect intestinal immune function and microbiota balance through the brain–gut axis, thereby exacerbating disease progression. Tongxie Yaofang, a classical formula for treating "liver stagnation and spleen deficiency" in TCM, has traditionally been used to regulate emotional disorders. Modern studies have confirmed that it can alleviate intestinal diseases by modulating the brain–gut axis. Tongxie Yaofang can regulate the levels of neurotransmitters such as 5-hydroxytryptamine (5-HT, serotonin) and calcitonin gene-related peptide (CGRP) in both the central and peripheral nervous systems, reduce visceral hypersensitivity, and alleviate abdominal pain, particularly in patients with IBS-D (diarrhea-predominant irritable bowel syndrome) and UC, thereby demonstrating the advantages of TCM in treating both the body and the mind (*69,70*). This therapeutic mechanism, which acts through modulation of the brain–gut axis, not only improves intestinal inflammatory responses but also helps relieve intestinal discomfort caused by psychological factors, highlighting the unique role of TCM in regulating the interaction among intestinal immunity, the nervous system, and emotional states. A detailed summary of the mechanism of action is provided in Figure 1.

4. The dual engines of the modernization of traditional Chinese medicine

Although classic TCM formulas have accumulated substantial clinical evidence for the treatment of IBD, and their multi-component and multi-target synergistic advantages have been scientifically elucidated through cellular and molecular mechanism studies, challenges remain in bridging traditional experience with modern precision medicine. On one hand, the complexity of formula components, poor water solubility of certain active ingredients, gastrointestinal degradation, and non-targeted tissue distribution lead to low bioavailability, limiting the maximal clinical efficacy (*71*). On the other hand, traditional research methods struggle to fully decipher the complex regulatory network of "multi-components—multi-targets—multi-pathways," hindering the precise identification of key active ingredients and their mechanisms of action (*72*). To overcome these challenges, two key technologies have become the "dual engines" driving the modernization of TCM formulas — nanomedicine through smart encapsulation and targeted regulation to address drug delivery issues, and multi-omics integration with network pharmacology to systematically clarify the core mechanisms. Together, they provide revolutionary solutions for the standardization, precision, and clinical translation of TCM formulas (Table 2).

4.1. Nanomedicine-based drug delivery systems

Nanomedicine, or nanodrug delivery systems (NDDS), is the core technology to solve drug delivery challenges in TCM. By precisely encapsulating and intelligently regulating active ingredients, NDDS can significantly enhance drug stability, targeting, and bioavailability while reducing systemic toxicity. In recent years, breakthrough advances have been made in the development of intestinal-targeted nanocarriers, especially in the areas of "environmentally responsive drug release" and "biological targeting modifications."

In the design of responsive carriers, pH-sensitive or reactive oxygen species (ROS)-responsive smart materials have become research hotspots due to the characteristics of the colonic inflammatory microenvironment. For example, low pH or high ROS-sensitive PLGA (Poly lactic-co-glycolic acid) nanoparticles enable selective drug release at the site of colonic inflammation, avoiding early degradation in the stomach and small intestine (*73*). Natural polysaccharides like hyaluronic acid and chitosan can specifically bind to the CD44 receptors overexpressed at the site of colonic inflammation, enhancing macrophage targeting (*74*). Additionally, by combining nanocarriers with microparticles or biomimetic coatings, multiple protective barriers can be constructed to further prevent enzymatic and acid-base degradation in the gastrointestinal tract, promoting efficient drug accumulation at the colonic site. In practical applications, the combination of nanotechnology and TCM ingredients with a "controlled release + targeted delivery" strategy effectively addresses the challenges of traditional formulations, significantly improving clinical applicability and safety. For example, berberine (berberine alkaloid) encapsulated in PLGA nanoparticles (BPL-NPs) can target and inhibit the IL-6/IL-6R signaling axis, reducing M1 macrophage infiltration in the colon, and significantly improving DSS-induced colitis (*75*). Similarly, composite nanoparticles prepared from the core components of Gegen Qinlian Decoction (berberine, puerarin, baicalin, and glycyrrhizic acid) activate the AMPK-PPARγ pathway to promote M2 macrophage polarization, enhancing drug accumulation at the ulcer site and balancing intestinal immunity (*46*). Additionally, nanotechnology can be used to encapsulate toxic TCM ingredients like triptolide, effectively reducing their accumulation in the kidneys and reproductive system, thereby minimizing side effects. For instance, silk fibroin nanoparticles (SFNP) of triptolide demonstrate enhanced anti-inflammatory and anti-tumor effects, significantly reducing liver and kidney toxicity in both cell and animal experiments (*76*). By loading triptolide and Atractylodes macrocephala into red blood cell membrane-coated liposomes (RBCm@Lip), immune cells can be "confused," prolonging the residence time of the drug in the blood and significantly reducing toxicity to normal tissues (*77*).

Overall, the synergistic effect of nanocarriers and TCM active ingredients significantly enhances

Table 2. Advanced technological strategies in traditional Chinese medicine–based IBD research: applications and key scientific/clinical problems addressed

Technological field	Specific technologies	Application examples in TCM-based IBD research	Key clinical problems addressed	Ref
Nanomedicine-based drug delivery systems	Plant-derived exosome-like nanovesicles (PELNs)	PELNs derived from grapes, ginger, and garlic for the treatment of colitis; loading with siRNA	Natural targeting: exploiting the intrinsic acid- and enzyme-resistant properties of plant vesicles to achieve colon-targeted delivery after oral administration; Safety: low immunogenicity, high biocompatibility, and absence of cytotoxicity; Therapeutic carrier–drug integration: the vesicles themselves exhibit intrinsic anti-inflammatory activity	(*84*)
	Smart stimulus-responsive nanocarriers	pH-responsive (targeting colonic pH), ROS-responsive (targeting the oxidative inflammatory microenvironment), and enzyme-responsive polymeric or liposomal encapsulation of triptolide and curcumin	Toxicity reduction and efficacy enhancement: drug release restricted to lesion sites (high ROS, specific pH), thereby markedly reducing the systemic toxicity of agents such as triptolide; Improved solubility and bioavailability of poorly soluble compounds (*e.g.*, curcumin)	(*73,76,77*)
	Colon-targeted microspheres/hydrogels	Chitosan-, pectin-, and alginate-based microspheres loaded with TCM formula extracts (*e.g.*, Baitouweng Decoction)	Prolonged retention: increased contact time and adhesion to the intestinal mucosa, enhanced local therapeutic effects, and reduced dosing frequency	(*74,83*)

drug delivery efficiency, bioavailability, and safety. Through chemical modification or biological coating of nanoparticles, characteristics can be conferred that respond to the intestinal environment (such as pH, ROS, and intestinal enzymes), further enabling targeted controlled release, maximizing the anti-inflammatory effect of TCM, and minimizing systemic exposure (*78*).

4.2. Multi-omics integration and network pharmacology

The integration of multi-omics and network pharmacology has provided a feasible path for systematically analyzing the "multi-component—multi-target—multi-pathway" mechanisms of TCM formulas (*72*). Multi-omics, based on high-throughput data, reveals the global coordination of TCM intervention on gene expression, metabolic networks, and gut microbiota ecology. Network pharmacology, through the construction of "drug—component—target—disease" association networks, identifies key active ingredients and core pathways, offering clear clues for experimental validation and translational research.

Several studies have demonstrated the value of this integrative approach: Wang *et al*. explored the mechanisms of Qing-Re-Hua-Shi Decoction (QRHSD) in treating UC, showing that QRHSD significantly modulates gut microbiota, increases beneficial Lactobacillus, inhibits pathogenic Morganella, and alters 507 metabolites, including bile acids and amino acid-related substances (*79*). Transcriptomic and network pharmacology analyses further revealed that QRHSD regulates gut homeostasis by downregulating key inflammatory signaling pathways such as PI3K/AKT, NF-κB, and MAPK. Additionally, correlation analysis demonstrated a close relationship between clinical indicators, microbiota changes, and metabolites, indicating that the formula restores gut homeostasis through a multi-layered mechanism. Another study used UPLC-QE-MS metabolomics and network pharmacology to analyze the mechanisms of Cynanchum wilfordii in treating UC, identifying 117 targets (such as JAK2, STAT3) and metabolic pathways related to inflammation through IL-17, HIF-1, and TNF signaling (*80*). Similarly, Wu *et al.* integrated metabolomics and network pharmacology to reveal the mechanism of Baicalin in experimental colitis, involving various metabolites and signaling pathways, including TNF, IL-6, and AKT1 (*81*). These studies have all employed the "multi-omics + network pharmacology" approach to systematically reveal the mechanisms of action of TCM formulas. They highlight the importance of multi-omics integration in uncovering the interactions between gut microbiota, metabolites, and signaling pathways, further deepening the understanding of TCM formulas' multi-target and multi-effect mechanisms.

Such studies underscore the core value of multi-omics integration in elucidating the interactions between gut microbiota, metabolites, and signaling pathways,

moving TCM mechanism research from "vague states" to "precise analysis." This provides scientific support for key ingredient selection, target validation, and new drug development, accelerating the modernization and clinical translation of TCM formulas.

5. Limitations

Several limitations of the current evidence base should be recognized. First, the currently available clinical evidence for classical TCM formulas in IBD is derived predominantly from ulcerative colitis rather than Crohn's disease, and many clinical studies are constrained by small sample sizes, single-center designs, geographically concentrated populations, and inadequate reporting of inclusion and exclusion criteria, thereby increasing the risk of selection bias and limiting the generalizability of the findings. Second, because IBD is a chronic relapsing disease, the relatively short follow-up period of many studies is insufficient to fully evaluate long-term efficacy, relapse prevention, and safety. Third, the compositional complexity and insufficient standardization of TCM formulas continue to hinder reproducibility and cross-study comparison. In addition, although TCM is characterized by multi-component and multi-target actions, many existing mechanistic studies remain focused on single pathways or targets, with limited validation of synergistic effects among different constituents. Furthermore, current animal models do not always align well with TCM syndrome differentiation, which may weaken the translational relevance of preclinical findings. Finally, the gut–brain axis, gut–lung axis, and nanodrug delivery systems discussed in this review should be regarded as promising but still emerging directions rather than clinically mature evidence. Future studies should emphasize better-designed multicenter trials, longer follow-up, improved standardization, and more integrative mechanistic validation.

6. Conclusion

IBD remains a major clinical challenge. Recent studies suggest that TCM is gradually moving from empirical use toward a more evidence-informed framework. Its core advantage lies in holistic and syndrome-oriented intervention strategies, which are inherently consistent with contemporary therapeutic goals, including mucosal healing, immune homeostasis, and microbial balance.

Mechanistic investigations have demonstrated that TCM does not merely exert general anti-inflammatory effects but regulates key pathological processes in IBD, including immune microenvironment remodeling, intestinal barrier repair, and microbiota–host metabolic interactions. These effects involve pathways such as NLRP3 inflammasome-mediated pyroptosis, macrophage polarization, Th17/Treg balance, and systemic regulatory axes such as the gut–lung and brain–gut axes. Meanwhile, emerging technologies, including nanomedicine-based delivery systems, multi-omics integration, and network pharmacology, are accelerating the modernization and mechanistic elucidation of TCM.

Future research should focus on precision-oriented and translational paradigms, integrating artificial intelligence, multi-omics data, and TCM syndrome classification to enable individualized therapy. Greater attention should also be given to long-term outcomes, safety evaluation, and integrative treatment strategies. Collectively, these efforts may establish TCM as a scientifically interpretable, mechanism-driven, and globally applicable therapeutic system for IBD.

Funding: This work was supported by the China Scholarship Council program (202506370114).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Kennedy MS, Chang EB. Emerging concepts and shifting paradigms for understanding the microbial basis of inflammatory bowel diseases. J Clin Invest. 2025; 135:e193969.
2. Gros B, Kaplan GG. Ulcerative Colitis in Adults: A Review. JAMA. 2023; 330:951-965.
3. Dolinger M, Torres J, Vermeire S. Crohn's disease. Lancet. 2024; 403:1177-1191.
4. Bruner LP, White AM, Proksell S. Inflammatory Bowel Disease. Prim Care. 2023; 50:411-427.
5. Hodson R. Inflammatory bowel disease. Nature. 2016; 540:S97.
6. Olfatifar M, Zali MR, Pourhoseingholi MA, Balaii H, Ghavami SB, Ivanchuk M, Ivanchuk P, Nazari SH, Shahrokh S, Sabour S, Khodakarim S, Aghdaei HA, Rohani P, Mehralian G. The emerging epidemic of inflammatory bowel disease in Asia and Iran by 2035: A modeling study. BMC Gastroenterol. 2021; 21:204.
7. Diez-Martin E, Hernandez-Suarez L, Muñoz-Villafranca C, Martin-Souto L, Astigarraga E, Ramirez-Garcia A, Barreda-Gómez G. Inflammatory Bowel Disease: A Comprehensive Analysis of Molecular Bases, Predictive Biomarkers, Diagnostic Methods, and Therapeutic Options. Int J Mol Sci. 2024; 25:7062.
8. Hracs L, Windsor JW, Gorospe J, *et al*. Global evolution of inflammatory bowel disease across epidemiologic stages. Nature. 2025; 642:458-466.
9. Ruan G, Sun Y, Yu Z, Bai X, Yang H, Qian J. Global, regional, and national burden of inflammatory bowel disease from 1990 to 2021: findings from the Global Burden of Disease 2021. Gastroenterol Rep (Oxf). 2025; 13:goaf082.
10. Stapersma L, van den Brink G, Szigethy EM, Escher JC, Utens E. Systematic review with meta-analysis: anxiety and depression in children and adolescents with inflammatory bowel disease. Aliment Pharmacol Ther. 2018; 48:496-506.
11. Fanizza J, Bencardino S, Allocca M, Furfaro F, Zilli A, Parigi TL, Fiorino G, Peyrin-Biroulet L, Danese S,

D'Amico F. Inflammatory Bowel Disease and Colorectal Cancer. Cancers (Basel). 2024; 16:2943.
12. Chang CW, Wong JM, Tung CC, Shih IL, Wang HY, Wei SC. Intestinal stricture in Crohn's disease. Intest Res. 2015; 13:19-26.
13. Yeshi K, Jamtsho T, Wangchuk P. Current treatments, emerging therapeutics, and natural remedies for inflammatory bowel disease. Molecules. 2024; 29:3954.
14. Burisch J, Claytor J, Hernandez I, Hou JK, Kaplan GG. The Cost of Inflammatory Bowel Disease Care: How to Make it Sustainable. Clin Gastroenterol Hepatol. 2025; 23:386-395.
15. Yang L, Luo H, Tan D, Zhang S, Zhong Z, Wang S, Vong CT, Wang Y. A recent update on the use of Chinese medicine in the treatment of inflammatory bowel disease. Phytomedicine. 2021; 92:153709.
16. Xiao Q, Mo D, Liu G, Xu Y, Chen Y, Lu C. Research progress on the diagnosis and treatment of intestinal fibrosis in Crohn's disease with traditional Chinese and Western medicine. Iran J Pharm Res. 2025; 24:e163617.
17. Shen Z, Zhou Q, Ni Y, He W, Shen H, Zhu L. Traditional Chinese medicine for mild-to-moderate ulcerative colitis: Protocol for a network meta-analysis of randomized controlled trials. Medicine (Baltimore). 2019; 98:e16881.
18. Sun YX, Wang X, Liao X, Guo J, Hou WB, Wang X, Liu JP, Liu ZL. An evidence mapping of systematic reviews and meta-analysis on traditional Chinese medicine for ulcerative colitis. BMC Complement Med Ther. 2021; 21:228.
19. Li Y, Zhu Z, He S, Tang J, Zhang Y, Yang Y, Dong Y, He L, Jia Y, Liu X. Shenling Baizhu Decoction treats ulcerative colitis of spleen-deficiency and dampness obstruction types by targeting 'gut microbiota and galactose metabolism-bone marrow' axis. J Ethnopharmacol. 2024; 335:118599.
20. Chen J, Shen B, Jiang Z. Traditional Chinese medicine prescription Shenling BaiZhu powder to treat ulcerative colitis: Clinical evidence and potential mechanisms. Front Pharmacol. 2022; 13:978558.
21. Chinese Association of Integrative Medicine. Expert consensus on the diagnosis and treatment of ulcerative colitis with integrated traditional Chinese and Western medicine. Zhong Xi Yi Jie He Za Zhi. 2023; 43:5-11. (in Chinese)
22. Chinese Association of Integrative Medicine. Guideline for the Diagnosis and Treatment of Ulcerative Colitis with Integrated Chinese and Western Medicine. Zhong Xi Yi Jie He Za Zhi. 2024; 44:1029-1035. (in Chinese)
23. Yang L, Song Y, Jin P, Liu Y, Wang Y, Qiao H, Huang Y. Shen-Ling-Bai-Zhu-San for ulcerative colitis: Protocol for a systematic review and meta-analysis. Medicine (Baltimore). 2018; 97:e12337.
24. Wu K, Luo J, Wu J, He T, Tan L, Li X. Meta-analysis of Shenling Baizhu Powder Combined with Aminosalicylic Acid Preparation in the Treatment of Ulcerative Colitis. Zhongguo Yao Fang. 2017; 28(36):5119-5122. (in Chinese)
25. Zhou T, Yang S, Wang J. Effects of Shenlingbaizhu powder combined with mesalazine in the treatment of ulcerative colitis and the influence of NLRP3 inflammasome. Zhongguo Lin Chuang Yao Li Xue Yu Zhi Liao Xue. 2018; 23:319-324. (in Chinese)
26. Li S, Hao X, Gong Y, Liu S, Niu W, Jia J, Tang Y. Effect of shenling baizhu powder on the serum TH1 cytokines of elderly patients with ulcerative colitis complicated by bloody purulent stool. Am J Transl Res. 2021; 13:9701-9707.
27. Weizhen Q, Weifeng Z. Effect of Shenlingbaizhu Powder Combined with Mesalamine on Inflammatory Factors and Immune Function in Children and Adolescents with Ulcerative Colitis. Yi Yao Dao Bao. 2019; 38:584-588. (in Chinese)
28. Jiao C, Zhang Q, Yang M, Ma J, Zhao X, Tang N, Dai M, Li Q, Jiang Z, Huang X, Zhang H, Sun L. Shenling Baizhu San ameliorates ulcerative colitis by regulating the gut microbiota and its tryptophan metabolites: A complementary medicine to mesalamine. J Ethnopharmacol. 2022; 291:115145.
29. Chen X, Cao J. Research Progress on the Treatment of Ulcerative Colitis with Shaoyao Decoction. Lin Chuang Ge Xing Hua Yi Xue. 2024; 3:1759-1766. (in Chinese)
30. Zhang S, Zhao L, Shen H, *et al.* International clinical practice guideline on the use of traditional Chinese medicine for ulcerative colitis by Board of Specialty Committee of Digestive System Disease of World Federation of Chinese Medicine Societies (2023). Phytother Res. 2024; 38:970-999.
31. Zhu Y, Liu Q. Safety and Efficacy of Shaoyao Decoction with Anti-Inflammatory Effect in the Treatment of Ulcerative Colitis: A Meta-Analysis. Iran J Public Health. 2025; 54:478-488.
32. Yao C, Li Y, Luo L, Feng P. Study on Mechanism of Shaoyao Decoction in Regulating Balance of Th17/Treg Cells to Improve Inflammatory Response of Ulcerative Colitis with Damp Heat Syndrome of Large Intestine. Shi Jie Ke Xue Ji Shu Zhong Yi Yao Xian Dai Hua. 2021; 23:2635-2642. (in Chinese)
33. Tan G, Sun J, Qu Y, Zhou F, Pei C, Feng T. Clinical observation of Shaoyao decoction and infliximab on moderate or severe ulcerative colitis. Shanxi Zhong Yi. 2020; 36:23-26. (in Chinese)
34. Xie Q, Yu W, He Y, Deng G, Yang H, Chen J, Liang J, Zeng X, Guo J, Ma R. Efficacy and safety of BaitouWeng decoction for ulcerative colitis: A meta-analysis of randomized and controlled trials. Medicine (Baltimore). 2024; 103:e38704.
35. Miao Z, Wang X, Zhang X, Zhang Y, Xu Y. Clinical efficacy assessment of new Baitouweng decoction in the treatment of active ulcerative colitis. Zhongguo Zhong Xi Yi Jie He Xiao Hua Za Zhi. 2021; 29:93-96. (in Chinese)
36. Ma Q, Li Y, Ding Y. Clinical observation of Baitouweng decoction modified combined with mesalazine sustained release tablets in the treatment of ulcerative colitis with hot toxicity and its effect on immune function and serum inflammatory factor in patients with ulcerative colitis. Hainan Yi Xue Yuan Xue Bao. 2019; 25:918-922. (in Chinese)
37. Liang G, Wu T, Sun H. Observation on the clinical efficacy of modified Baitouweng decoction retention enema in the treatment of ulcerative colitis of damp-heat accumulation type under the theory of chronic disease entering collaterals. Guo Ji Lin Chuang Yan Jiu. 2024; 8:17-23. (in Chinese)
38. Yu Y, Wang Q, Chen Y, Li C, Zhang M. Regulatory effect of Tongxie Yaofang on immune function and intestinal mucosal barrier in patients with ulcerative colitis. Zhongguo Zhong Xi Yi Jie He Xiao Hua Za Zhi. 2020; 28:858-862. (in Chinese)
39. Shang Z, Zhou L, Liu Y, Yang X, Yao D, Zhai F, Liu B, Chen Y, Zhu X, Liu D, Yi B. Tongxie Yaofang attenuates ulcerative colitis by modulating gut microbiota and

IL-10RA/NF-κB-mediated macrophage polarization. Phytomedicine. 2025; 148:157459.
40. Du C. Study on the Overall Effect of Gegen Qinlian Decoction Combined with Mesalazine in the Treatment of Ulcerative Colitis. Medical Tribune. 2023; 5:42-44. (in Chinese).
41. Chen T, Wang Z, Peng Y, Wang Q, Pei Y, Yang Q, Yang W. Clinical effects of Supplemented Gegen Qinlian Decoction combined with acupuncture on patients with ulcerative colitis of Large Intestinal Dampness-heat Pattern. Zhong Cheng Yao. 2025; 47:453-457. (in Chinese)
42. Chen Y, Li R, Yuan Y, He H, Yang W, Yuan C, Xiao J, Qin Y, Lian M, Yang Z. Efficacy and safety of Gegen Qinlian decoction combined with mesalazine for the treatment of UC: A meta-analysis and systematic review. Medicine (Baltimore). 2025; 104:e42266.
43. Liu C, Wang J, Yang Y, Liu X, Zhu Y, Zou J, Peng S, Le TH, Chen Y, Zhao S, He B, Mi Q, Zhang X, Du Q. Ginsenoside Rd ameliorates colitis by inducing p62-driven mitophagy-mediated NLRP3 inflammasome inactivation in mice. Biochem Pharmacol. 2018; 155:366-379.
44. Wei YY, Fan YM, Ga Y, Zhang YN, Han JC, Hao ZH. Shaoyao decoction attenuates DSS-induced ulcerative colitis, macrophage and NLRP3 inflammasome activation through the MKP1/NF-κB pathway. Phytomedicine. 2021; 92:153743.
45. Liu L, Miao ZW, Wei YZ, Bu S, Gu X, Xu Y, Shan ZW. New Baitouweng Decoction alleviated DSS-induced colitis through the FXR/NLRP3 signaling pathway by regulating gut microbiota and bile acids. Gastroenterol Rep (Oxf). 2025; 13:goaf055.
46. Li J, Cao J, Zhang Z, Wang S, Zhu M, Yang L, Ouyang W, Ma C. Mechanism study on the treatment of ulcerative colitis by Gegen Qinlian nano-preparation through promoting M2 macrophage polarization. Front Mol Biosci. 2025; 12:1580874.
47. Yang Z, Lin S, Feng W, Liu Y, Song Z, Pan G, Zhang Y, Dai X, Ding X, Chen L, Wang Y. A potential therapeutic target in traditional Chinese medicine for ulcerative colitis: Macrophage polarization. Front Pharmacol. 2022; 13:999179.
48. Hu Y, Yang Y, Li Y, Zhang Q, Zhang W, Jia J, Han Z, Wang J. Th17/Treg imbalance in inflammatory bowel disease: immunological mechanisms and microbiota-driven regulation. Front Immunol. 2025; 16:1651063.
49. Chen L, Ruan G, Cheng Y, Yi A, Chen D, Wei Y. The role of Th17 cells in inflammatory bowel disease and the research progress. Front Immunol. 2022; 13:1055914.
50. Zhao D, Ge A, Yan C, Liu X, Yang K, Yan Y, Hao M, Chen J, Daga P, Dai CC, Li C, Cao H. T helper cell 17/regulatory T cell balance regulates ulcerative colitis and the therapeutic role of natural plant components: a review. Front Med (Lausanne). 2024; 11:1502849.
51. Villablanca EJ, Selin K, Hedin CRH. Mechanisms of mucosal healing: treating inflammatory bowel disease without immunosuppression? Nat Rev Gastroenterol Hepatol. 2022; 19:493-507.
52. Luo Z, Li Z, Liang Z, Wang L, He G, Wang D, Shen L, Wang Z, Ma X, Geng F, Wang H, Liu W, Liu H, Li B. Berberine increases stromal production of Wnt molecules and activates Lgr5(+) stem cells to promote epithelial restitution in experimental colitis. BMC Biol. 2022; 20:287.
53. Kay SK, Harrington HA, Shepherd S, Brennan K, Dale T, Osborne JM, Gavaghan DJ, Byrne HM. The role of the Hes1 crosstalk hub in Notch-Wnt interactions of the intestinal crypt. PLoS Comput Biol. 2017; 13:e1005400.
54. Meir M, Kannapin F, Diefenbacher M, Ghoreishi Y, Kollmann C, Flemming S, Germer CT, Waschke J, Leven P, Schneider R, Wehner S, Burkard N, Schlegel N. Intestinal Epithelial Barrier Maturation by Enteric Glial Cells Is GDNF-Dependent. Int J Mol Sci. 2021; 22:1887.
55. Santhosh S, Zanoletti L, Stamp LA, Hao MM, Matteoli G. From diversity to disease: unravelling the role of enteric glial cells. Front Immunol. 2024; 15:1408744.
56. Zhou M, Chen Y, Jin W, Li P, Hu J, Guo X. Traditional Chinese Medicine: A Promising Treatment Option for Intestinal Fibrosis. Am J Chin Med. 2024; 52:2107-2129.
57. Li Y, Xu F, Fang Y, Cui Y, Zhu Z, Wu Y, Tong Y, Hu J, Zhu L, Shen H. Inflammation-fibrosis interplay in inflammatory bowel disease: mechanisms, progression, and therapeutic strategies. Front Pharmacol. 2025; 16:1530797.
58. Li P, Xu R, Chen H, Cao S, Xie X, Xu B, Chen Y, Xie X, Yang C, Wang Q, Luo X, Zhou L. The promoting effect of high-fat diet-induced trained immunity on ulcerative colitis and the mechanism of the therapeutic effect of Dahuang Mudan Decoction. Phytomedicine. 2025; 143:156893.
59. Zhang J, Lin B, Zhang Y, Hu X, Liu T, Liu EH, Liu S. Baitouweng decoction alleviates ulcerative colitis by regulating tryptophan metabolism through DOPA decarboxylase promotion. Front Pharmacol. 2024; 15:1423307.
60. Jing W, Dong S, Xu Y, Liu J, Ren J, Liu X, Zhu M, Zhang M, Shi H, Li N, Xia P, Lu H, Wang S. Gut microbiota-derived tryptophan metabolites regulated by Wuji Wan to attenuate colitis through AhR signaling activation. Acta Pharm Sin B. 2025; 15:205-223.
61. Huang J, Zhang J, Wang F, Tang X. Modified Gegen Qinlian Decoction modulated the gut microbiome and bile acid metabolism and restored the function of goblet cells in a mouse model of ulcerative colitis. Front Immunol. 2024; 15:1445838.
62. Yang J, Huang Q, Long J, Li J. Microbiota and inflammatory bowel disease: the dual effect mechanism of polysaccharide therapy. Front Immunol. 2025; 16:1666866.
63. Günther C, Rothhammer V, Karow M, Neurath M, Winner B. The Gut-Brain Axis in Inflammatory Bowel Disease-Current and Future Perspectives. Int J Mol Sci. 2021; 22:8870.
64. Ziaka M, Exadaktylos A. Gut-derived immune cells and the gut-lung axis in ARDS. Crit Care. 2024; 28:220.
65. Wang H, Wang Y. What Makes the Gut-Lung Axis Working? From the Perspective of Microbiota and Traditional Chinese Medicine. Can J Infect Dis Med Microbiol. 2024; 2024:8640014.
66. Wurbel MA, McIntire MG, Dwyer P, Fiebiger E. CCL25/CCR9 interactions regulate large intestinal inflammation in a murine model of acute colitis. PLoS One. 2011; 6:e16442.
67. Yu Q, Sun S, Han T, Li H, Yao F, Zong D, Li Z. Investigating the Therapeutic Mechanisms of Shen-Ling-Bai-Zhu-San in Type 2 Diabetes and Ulcerative Colitis Comorbidity: A Network Pharmacology and Molecular Simulation Approach. Pharmaceuticals (Basel). 2025; 18:1516.
68. Dong Y, He L, Zhu Z, Yang F, Ma Q, Zhang Y, Zhang X, Liu X. The mechanism of gut-lung axis in pulmonary fibrosis. Front Cell Infect Microbiol. 2024; 14:1258246.
69. Liang SB, Cao HJ, Kong LY, Wei JL, Robinson N, Yang SH, Zhu SJ, Li YQ, Fei YT, Han M, Liu JP. Systematic review and meta-analysis of Chinese herbal formula

Tongxie Yaofang for diarrhea-predominant irritable bowel syndrome: Evidence for clinical practice and future trials. Front Pharmacol. 2022; 13:904657.
70. Liang SB, Cheng HJ, Zhang QY, *et al.* Chinese herbal formula Tongxie Yaofang granules for diarrhoea-predominant irritable bowel syndrome: a randomised, double-blind, placebo-controlled, phase II trial. BMJ Open. 2025; 15:e088410.
71. Zhang YB, Wang JF, Wang MX, Peng J, Kong XD, Tian J. Nano-based drug delivery systems for active ingredients from traditional Chinese medicine: Harnessing the power of nanotechnology. Front Pharmacol. 2024; 15:1405252.
72. Zhao W, Wang B, Li S. Network pharmacology for traditional Chinese medicine in era of artificial intelligence. Chin Herb Med. 2024; 16:558-560.
73. Wang J, Yao M, Zou J, Ding W, Sun M, Zhuge Y, Gao F. pH-Sensitive Nanoparticles for Colonic Delivery Anti-miR-301a in Mouse Models of Inflammatory Bowel Diseases. Nanomaterials (Basel). 2023; 13:2797.
74. Hlaing SP, Cao J, Lee J, Kim J, Saparbayeva A, Kwak D, Kim H, Hwang S, Yun H, Moon HR, Jung Y, Yoo JW. Hyaluronic Acid-Conjugated PLGA Nanoparticles Alleviate Ulcerative Colitis *via* CD44-Mediated Dual Targeting to Inflamed Colitis Tissue and Macrophages. Pharmaceutics. 2022; 14:2118.
75. Liu C, Gong Q, Liu W, Zhao Y, Yan X, Yang T. Berberine-loaded PLGA nanoparticles alleviate ulcerative colitis by targeting IL-6/IL-6R axis. J Transl Med. 2024; 22:963.
76. Mei Z, Li X, Wu Q, Hu S, Yang X. The research on the anti-inflammatory activity and hepatotoxicity of triptolide-loaded solid lipid nanoparticle. Pharmacol Res. 2005; 51:345-351.
77. Yan T, Wang Y, Hou Z, Song P, Wang P, Li J, Zheng J, Lv D. A comprehensive overview of triptolide utilizing nanotechnology and its potential applications in prostate diseases. Front Pharmacol. 2025; 16:1592066.
78. Chen M, Lan H, Jin K, Chen Y. Responsive nanosystems for targeted therapy of ulcerative colitis: Current practices and future perspectives. Drug Deliv. 2023; 30:2219427.
79. Chen W, Xu L, Wang L, Shan YN, Li Y, Zhu JS. Qing-Re-Hua-Shi Decoction ameliorates DSS-induced colitis by modulating multiple signaling pathways and remodeling the gut microbiota and metabolite profile. Front Cell Infect Microbiol. 2025; 15:1541289.
80. Yu G, Cheng W, Tu X, Zhang M, Li H, Nie J. Therapeutic mechanism of Cynanchum wilfordii for ulcerative colitis: an analysis using UPLC-QE-MS, network pharmacology and metabolomics. Nan Fang Yi Ke Da Xue Xue Bao. 2024; 44:1485-1496. (in Chinese)
81. Wu Q, Wu X, Wang M, Liu K, Li Y, Ruan X, Qian L, Meng L, Sun Z, Zhu L, Wu J, Mu G. Therapeutic Mechanism of Baicalin in Experimental Colitis Analyzed Using Network Pharmacology and Metabolomics. Drug Des Devel Ther. 2023; 17:1007-1024.
82. Shu X, Cao Y, Wu Y, Chen M, Zhao W, Ji G, Zhang L. Gegen-Qinlian decoction alleviates metabolic dysfunction-associated steatohepatitis by modulating the microbiota-bile acid axis in mice. J Ethnopharmacol. 2025; 347:119719.
83. Liu H, Jin X, Liu S, Liu X, Pei X, Sun K, Li M, Wang P, Chang Y, Wang T, Wang B, Yu XA. Recent advances in self-targeting natural product-based nanomedicines. J Nanobiotechnology. 2025; 23:31.
84. Fang X, Feng J, Zhu X, Feng D, Zheng L. Plant-derived vesicle-like nanoparticles: A new tool for inflammatory bowel disease and colitis-associated cancer treatment. Mol Ther. 2024; 32:890-909.
85. Fang Y, Min S, Wu Y, Xu F, Chen H, Li Y, Lu Y, Hu J, Zhu L, Shen H. Integration of multi-omics and network pharmacology analysis reveals the mechanism of Qingchang Huashi Jianpi Bushen formula in repairing the epithelial barrier of ulcerative colitis. J Inflamm Res. 2025; 18:6167-6189.

Received February 23, 2026; Revised April 16, 2026; Accepted April 18, 2026.

[§]These authors contributed equally to this work.

**Address correspondence to:*
Lianwen Yuan, Department of General Surgery, The Second Xiangya Hospital, Central South University, 139 Renmin Middle Road, Changsha City, Hunan Province, China.
E-mail: yuanlianwen@csu.edu.cn

Released online in J-STAGE as advance publication April 24, 2026.

Review

DOI: 10.5582/ddt.2026.01012

Modern research on formulated granules of traditional Chinese medicines and Japanese Kampo medicines: A narrative review

Wenyuan Li

School of Pharmacy, Qinghai University, Xining, Qinghai, China.

SUMMARY: The modernization of traditional medicine follows diverse pathways. This paper aims to provide a narrative review of the manufacturing processes, research on active ingredients, and models for the clinical use of formulated granules of traditional Chinese medicines (TCM) and Japanese Kampo medicines in the context of their modernization by comparing advances in research on the two and their challenges. By searching databases such as PubMed, Web of Science, and China National Knowledge Infrastructure (CNKI), this review included relevant literature published up to 2026. After screening, a total of 48 relevant articles were included in the review. In TCM, formulated granules evolved from the need to modernize herbal medicines through multi-stage extraction and granulation. This enhances convenience while preserving medicinal properties. TCM enables flexible syndrome differentiation and it is working to industrially produce and standardize herbal medicines while refining quality standards. Japanese Kampo, based on fixed classical formulations and covered by National Health Insurance, has developed a highly standardized, evidence-based model. Research in China broadly investigates numerous formulations using omics technologies to explore their material basis and network pharmacology. Research in Japan focuses on in-depth analysis of "pathways to efficacious formulations" for select classic formulations. Faced with the challenge of providing modern evidence, granules can produce real-world data on TCM. This comparative analysis offers insights into the development of quality granules and their use internationally and it explores prospects for Sino-Japanese collaboration in complementary manufacturing processes and fundamental research.

Keywords: formulated granules of traditional Chinese medicine, Kampo medicine, modernization

1. Introduction

Traditional Chinese medicine (TCM) represents one of the most remarkable achievements of Chinese civilization, serving as a medical science with a heritage spanning millennia. Developed from empirical knowledge amassed in everyday life, TCM has been continuously refined through generations in the battle against diseases. It has made significant contributions to combating the SARS-CoV-2 virus and treating COVID-19 (*1*). In 2015, Tu Youyou's Nobel Prize elevated artemisinin to a new global status as "a gift from TCM to the world" (*2*). At present, TCM is extensively practiced in 183 countries and regions globally, with 103 countries having authorized the use of acupuncture therapy (*3*).

The modernization of TCM constitutes a central challenge to its preservation and innovation. In the process of promoting the modernization of the TCM industry, China has developed formulated granules as a significant dosage form. This innovation is designed to tackle the inconvenience associated with traditional decoction preparation and the challenges involved in quality control, while keeping pace with advances in modern healthcare. The path of development of formulated granules in TCM, as a representative foray into the modernization of traditional medicine, has been extensively studied and recognized for its innovative approach and significant impact. Japan, with its longstanding practice of TCM, has evolved a highly systematized form of traditional medicine known as Kampo. This system has been developed over centuries, with significant influence from ancient Chinese medical texts such as the *Shang Han Za Bing Lun*, and it has been integrated into the modern Japanese healthcare system. Kampo, which has fixed prescriptions and standardized extraction processes, has been integrated into Japan's National Health Insurance system, signifying its acceptance as a standardized form of traditional medicine. A systematic comparison of the similarities and differences between formulated granules in TCM and Kampo medicine in terms of manufacturing

processes, research efforts, and clinical uses will not only elucidate the unique features and challenges of TCM's development in China but also offer valuable insights from an international perspective. This review focuses on formulated granules in TCM, contrasting them with Kampo medicine to analyze differences in technological systems, active ingredient research, and regulations on clinical use. Formulated granules need to be standardized and their quality controlled, but at the same time the essence of TCM needs to be maintained while adopting modern scientific methods. This review aims to provide references for the continuous optimization and scientific advancement of traditional medicines.

2. Methods

To prepare this narrative review, the author conducted a systematic search of PubMed, Web of Science, Google Scholar, and China National Knowledge Infrastructure (CNKI). Search terms included "formulated granules of traditional Chinese medicine" and "Kampo medicine." Studies addressing process systems, active ingredient research, and clinical uses were included. After removing duplicates, conducting an initial screening of titles and abstracts, and conducting a secondary screening of full-text articles, a total of 48 studies were ultimately included in this review. As this is a narrative review aimed at providing a general overview of representative advances in the field, no assessment of study quality or meta-analysis was performed.

3. Historical context and industrialization

3.1. Emergence and development of formulated granules in TCM

The development of formulated granules in TCM signifies a major step in the modernization of Chinese herbal decoctions. Its conceptualization and implementation commenced in the late 20th century, and it entered a pilot exploratory phase in the early 21st century following the promulgation of the Interim Regulations on the Management of Formulated Granules of Traditional Chinese Medicine. Following more than two decades of development, the pilot phase concluded in 2021, marking the advent of a new era guided by national standards. This marked the transition from an "exploratory product" to a "standardized pharmaceutical" (*4*). This development was spurred by the urgent demand for greater convenience, consistency, and quality control in the clinical use of TCM. The primary goal is to offer, *via* industrialized production, a modern TCM in a dosage form that matches traditional decoctions in efficacy while being more convenient to dispense and administer. A core set of processes has been established, encompassing extraction of water, concentration, drying, and granulation, though special processing requirements for certain herbs (*e.g.*, pre-decoction or post-addition) are also addressed. Efforts are concentrated on developing a holistic quality assurance framework that encompasses the entire lifecycle of TCM products, from the selection of raw herbs and processing of herbal materials to quality control for the final product (*5*). As of this writing, provincial drug standards directories include 518 varieties, 350 of which are formulated granules that are recognized by national drug standards (*6*).

3.2. Revival and modernization of Japanese Kampo

The developmental trajectory of Japanese Kampo differs significantly from that of TCM. Once TCM was introduced to Japan, it underwent a process of local adaptation, resulting in the creation of Kampo medicine. However, Kampo was marginalized during the Meiji Restoration. A revival of Kampo took place in the mid-20th century, amid growing reflections on the side effects of chemical drugs. A pivotal moment in its modernization was the inclusion of Kampo extract preparations in the National Health Insurance system in 1967, which marked the first approval of Kampo medicines for reimbursement. This significantly expedited their industrial production, standardization, and clinical use (*7*). According to data from 2015, Kampo medicines constituted approximately 2.4% of Japan's total pharmaceutical sales, with the majority being prescription-only dry extract formulations, as reported by industry sources. Currently, 148 traditional medicine formulations that meet ethical standards, similar to Japanese Kampo medicine, are covered by national health insurance. While official certification standards for 294 over-the-counter traditional medicine products have been formally established, the number of such products has been steadily increasing, reflecting a shift towards greater accessibility and consumer autonomy in healthcare (*8*). Japan has established an industrial model centered on fixed classical formulations, primarily derived from texts such as the *Shang Han Za Bing Lun and Jin Gui Yao Lue* (Table 1), and featuring concentrated production by a few large enterprises, such as Tsumura Pharmaceutical. The national pharmacopoeia, industry GMP standards and stringent requirements for equivalence with standard decoctions together constitute a mature quality regulatory framework (*7*). Since 2001, Kampo medicine education has also been incorporated into Japan's core medical curriculum (*9*).

4. Comparison of process technology systems

4.1. Process characteristics of formulated granules in TCM

The core process of formulated granules in TCM involves the production of single-ingredient concentrated extracts from traditional Chinese herbs. Key technical

Table 1. Comparison of commonly used TCM prescriptions and Kampo prescriptions

No.	TCM Prescriptions	Kampo Prescriptions (*47*)	No.	TCM Prescriptions	Kampo Prescriptions (*47*)
1	Liujunzitang	Rikkunshito	21	Chaihuguizhiganjiangtang	Saikokeishikankyoto
2	Buzhongyiqitang/Buzhongyiqikeli (*46*)	Hochuekkito	22	Xiaojianzhongtang/ Xiaojianzhongkeli (*46*)	Shokenchuto
3	Gegentang/Gegentangkeli (*46*)	Kakkonto	23	Dangguisinijiawei	Tokishigyakukagoshuyushokyoto
4	Guizhifulingwan (*46*)	Keishibukuryogan	24	Chaihuguizhitang	Saikokeishito
5	Jiaweixiaoyaowan (*46*)	Kamishoyosan	25	Maimendongtang	Bakumondoto
6	Baweidihuangwan/Shenqiwan	Hachimijiogan	26	Sinisan	Shigyakusan
7	Jishengshenqiwan (*46*)	Goshajinkigan	27	Xiaoqinglongtang/ Xiaoqinglongkeli (*46*)	Shoseiryuto
8	Shaoyaogancaotang	Shakuyakukanzoto			
9	Wulingsan (*46*)	Goreisan	28	Mahuangfuzixixintang	Maobushisaishinto
10	Dangguishaoyaosan	Tokishakuyakusan	29	Guizhijiashaoyaotang	Keishikashakuyakuto
11	Banxiahouputang	Hangekobokuto	30	Xiaochaihutang/Xiaochaihukeli (*46*)	Shosaikoto
12	Yigansan	Yokukansan			
13	Lingguizhugantang	Ryokeijutsukanto	31	Wujisan	Goshakusan (Ojeoksan)
14	Dajianzhongtang	Daikenchuto	32	Zhulingtang	Choreito
15	Yigansanjiachenpibanxia	Yokukansankachinpihange	33	Guizhifulingwanjiawei	Keishibukuryogankayokuinin
16	Renshenyangrongtang/ Renshenyangrongwan (*46*)	Ninjinyoeito	34	Siwutangjiawei	Shichimitsukokato
			35	Maxingshigantang	Makyoyokukanto
17	Shiquandabutang/ Shiquandabuwan (*46*)	Juzentaihoto	36	Taohechengqitang	Tokakujokito
			37	Banxiaxiexintang	Hangeshashinto
18	Chaihujialonggumuli	Saikokaryukotsuboreito	38	Yinchenhaotang	Inchinkoto
19	Shujinghuoxuetang	Sokeikakketsuto	39	Fangjihuangqitang	Boiogito
20	Guipitang/Guipikeli (*46*)	Kamikihito	40	Mazirenwan	Mashiningan

features include: (1) Adaptive extraction: For materials with distinct properties, such as minerals, toxic components, precious herbs, and volatile constituents, strategies like pre-decoction, post-addition, separate extraction, and wrapped decoction are traditional techniques used in TCM to ensure the effective display of medicinal properties and to mitigate toxicity. Mineral and shell-based materials like *Gypsum Fibrosum* and *Ostreae Concha* require pre-decoction. Chinese medicinal materials that exhibit thermal instability, heat-induced volatility, or pose toxicity risks due to prolonged boiling, such as *Menthae Haplocalycis Herba* and *Foeniculi Fructus*, are added later in the extraction process. Chinese medicinal materials that may cause decoction viscosity or turbidity, like *Typhae Pollen* and *Plantaginis Semen*, require enclosed decoction. Chinese medicinal materials with a high viscosity (*e.g.*, *Asini Corii Colla* and *Testudinis Carapacis et Plastri Colla*) require melting to minimize interference with other materials during extraction. Certain precious Chinese medicinal materials, such as *Ginseng Radix et Rhizoma* and *Panacis Quinquefolii Radix*, need to be decocted separately to maintain their activity. Highly valuable materials, such as *Bovis Calculus* and *Moschus*, are typically pulverized and taken as an infusion with a decoction of other materials; (2) Full-process quality control: Emphasizes the correlation of characteristic spectra or fingerprint spectra across "decoction pieces-intermediates-finished products" to demonstrate consistency with the material basis of traditional decoctions. The Technical Requirements for Quality Control and Setting Standards for Formulated Granules of Traditional Chinese Medicines stipulate that the specific or fingerprint spectra of Chinese medicinal materials , decoction pieces, intermediates, and granules must be correlated. This standard emphasizes that extraction should be based on traditional decoctions, ensuring consistency in extraction yield, key components, and characteristic spectra. Standardized decoctions are usually analyzed alongside extracts or granules to pinpoint factors influencing the consistency of quality during extraction (*10*); and (3) Technological diversification: Widespread adoption of dry granulation, spray-drying granulation, and wet granulation (*e.g.*, extrusion granulation, fluidized bed granulation, and high-shear wet granulation) (*11*). The underlying process logic involves flexibly using formulated components in industrial production in order to strike a balance between production efficiency and the inherent complexity and individualization of traditional Chinese clinical prescriptions.

4.2. Characteristics of Japanese Kampo manufacturing processes

Japanese Kampo technology centers on the core principles of "restoration and standardization" and seeks to modernize and standardize the production of traditional formulations. Its key characteristics include: (1) Fixed formulations and raw materials: Utilizing pre-formulated mixtures of "crude drugs" as starting materials; (2) Standardized decoction and drying: Preparing standardized decoction extracts under strictly controlled conditions (time, temperature, and solvent) that are primarily processed into granules *via* spray-drying technology. This decoction essentially originates from traditional decoction methods, in which raw herbs

are boiled to extract their soluble medicinal components; and (3) Stringent equivalence criteria: Finished products are required to exhibit "equivalence" with standard decoctions based on an analysis of the content of target components and/or fingerprint patterns. Production processes strictly comply with pharmaceutical GMP standards and are supported by established quality control systems. This approach embodies the principle of "fixed ancient formulations, precise replication," guaranteeing a high level of product consistency and stability (*8*).

To safeguard the quality of Kampo formulations, Japan's Ministry of Health, Labor, and Welfare issued the Guide to Data Requirements for Ethical Kampo Formulations in 1985. This requires systematic comparisons between Kampo extract formulations and standard decoctions to determine whether their quality control markers are equivalent. Related studies are required to provide data from at least three batches, and each batch needs to be tested at least three times (*12*).

In addition, Japan developed a system of abbreviated English notation for Kampo preparations to promote the international spread of Kampo medicine. This system covers all 298 Kampo preparations listed in the 2013 New Guidelines for Over-the-Counter Kampo Preparations, including non-prescription Kampo extract preparations (*13*).

In summary, the strength of Chinese manufacturing processes lies in their high degree of flexibility, which effectively aligns with the clinical practice of TCM syndrome differentiation and treatment, thereby meeting individual therapeutic needs. The complexity of those processes hampers the achievement of a high level of consistency across different manufacturers and batches, since quality standards are still being standardized. Japanese manufacturing processes are renowned for their rigorous standardization, which ensures consistent quality and facilitates systematic pharmacological and clinical research. A drawback is the use of fixed formulations, which cannot be adjusted based on syndrome differentiation, leading to limited flexibility.

5. Research on active components and pharmacologically active substances

5.1. Research on active components in formulated granules

Research on formulated granules in TCM primarily focuses on two directions: (1) Validation studies on the "equivalence" between formulated granules and traditional decoctions. Recent studies have systematically compared the chemical profiles of traditional decoctions and formulated granules using modern analytical techniques such as high-resolution mass spectrometry, revealing that formulated granules may offer advantages in terms of convenience and effectiveness. Evidence of their consistent quality is being provided. For example, a quantitative analysis of 64 components, including glycyrrhizic acid and naringin, in Erchen Decoction Granules using UPLC-Q-Exactive Orbitrap MS corroborated the chemical equivalence of those components (*14*); and (2) Research on the pharmacological mechanisms and action pathways of formulated granules. This research direction comprehensively utilizes network pharmacology, proteomics, metabolomics, and animal models to elucidate the multi-target mechanisms by which formulated granules are effective in treating complex diseases. Below are a few examples.

Biyuan Tongqiao Granules: Ultra-high-performance liquid chromatography-electrospray ionization-quadrupole-linear ion trap-mass spectrometry (UHPLC-ESI-QE-Orbitrap-MS) was used to identify 58 compounds in that medicine (*15*). By combining an ovalbumin (OVA)-induced allergic rhinitis mouse model with proteomics analysis, that study revealed that it acted to regulate the PI3K/AKT and STAT3/MAPK pathways to inhibit the expression of E-selectin, VCAM-1, and ICAM-1, thereby alleviating allergic rhinitis.

Jingfang Granules: In a bleomycin-induced acute lung injury model, Jingfang Granules were confirmed to have a lung-protective effect by downregulating the PI3K/Akt/mTOR signaling pathway and modulating glycolysis/gluconeogenesis and pyruvate metabolism (*16*).

Qishen Granules: In an adriamycin-induced cardiac toxicity model, Qishen Granules were found to alleviate oxidative damage, protect mitochondrial function by activating the SIRT3/Ac-SOD2 signaling pathway, and thereby improve doxorubicin-induced cardiotoxicity (*17*).

Jieduhuayu Granules: Using a combined metabolomics and proteomics analysis approach, a study evaluated changes in proteins and metabolites in liver tissue from rats with alcoholic fatty liver disease before and after treatment with Detoxifying and Blood-Stasis-Resolving Granules (*18*). The findings suggest that the mechanism may involve activation of the PI3K-AKT pathway and inhibition of hepatic glycolysis, thereby alleviating hepatic inflammatory stress.

Additionally, studies on the correlation between spectra and effects are ongoing. For example, high-performance liquid chromatography fingerprinting technology was used to systematically analyze the relationship between the active component groups (total flavonoids, total saponins, total alkaloids, *etc.*) in Yinyang Tongnao Granules and the pharmacological efficacy indicators of chronic intermittent hypoxic-ischemic reperfusion injury in rats, providing elucidation of the material basis underlying that medicine's holistic effects (*19*).

5.2. Active component research in Japanese Kampo

Japanese research adopts a "fewer but more refined" strategy, concentrating on classical fixed formulations covered by National Health Insurance to conduct in-depth and systematic analyses of their pharmacological mechanisms. Studies underscore the importance of incorporating chemical analysis, cellular assays, animal experimentation, and clinical observations to comprehensively understand the integrated pathways of compound formulations.

For example, Choreito has been found to alleviate detrusor overactivity and bladder pain symptoms induced by tranilast in a rat model, mimicking interstitial cystitis/painful bladder syndrome, as evinced by a recent study (*20*). This aligns with the broader context of interstitial cystitis treatment, which includes a range of approaches such as drug therapy and surgical interventions, as discussed in various medical sources. An exploratory study with an open-label, single-arm design further demonstrated significant improvement in bladder pain scores following treatment with this formulation, indicating its therapeutic efficacy for bladder pain (*21*). This type of study design is often used in early clinical trials, such as Phase I and II, to explore safety and efficacy, as well as to provide preliminary data on the treatment's potential benefits.

Yokukansan is clinically recognized as beneficial for dementia patients. Recent studies using tauopathy cell models have demonstrated that a specific formulation can decrease levels of phosphorylated and oligomeric tau by suppressing GSK3β kinase activity and promoting autophagy. This mechanism holds promise for intervening in the pathology of Alzheimer's disease, as evinced by the success of GSK-3β inhibitors in reducing phosphorylated Tau protein levels (*22*).

Retrospective clinical analysis combined with molecular mechanism studies of Boiogito have revealed that this formulation promotes LRRC8A transport to the plasma membrane and activates the VSOR chloride channel, thereby inducing chloride ion release and water excretion (*23*). This provides scientific evidence of its clinical efficacy.

These studies exemplify Japanese Kampo medicine's emphasis on translational research—from cellular and animal model studies to clinical observations—in deciphering "pathways to efficacious formulations," establishing a research avenue where basic and clinical evidence merge.

In summary, Chinese research covers a broad range of medicines and involves cutting-edge techniques, actively exploring the material basis and mechanisms of numerous formulated granules on the market. However, systematic studies on individual formulations need to go into further depth and clinical translation needs to be enhanced further. Japanese research, and particularly that in the field of Kampo medicine, excels in depth and clinical integration, building robust evidence chains from molecular mechanisms to clinical efficacy through the use of modern research methodologies and fixed formulations. These differing strategies reflect distinct industrial models: China has a vast formulation system in the early stages of industrialization and a vast need, while Japan focuses on in-depth development of a limited number of formulations for industrial production.

6. Clinical use system and regulatory framework

6.1. Formulated granules of TCM

The clinical use of formulated granules of TCM is centered on syndrome differentiation and treatment in TCM, enabling physicians to tailor prescriptions to an individual patient's condition. This approach ensures personalized treatment and has been found to be effective in various clinical settings, as evinced by studies demonstrating the efficacy and convenience of granules. TCM's regulatory framework has evolved from trial local standards to unified national standards (*6*). As a further example, a prospective cohort study demonstrated that Jingyin Granules can enhance nucleic acid conversion rates, shorten conversion time, and reduce the duration of hospitalization for patients with mild COVID-19 (*24*). A recent randomized controlled trial, following the CONSORT guidelines and utilizing statistical software for data analysis, demonstrated that Qingjin Yiqi Granules significantly alleviated dyspnea and fatigue symptoms in patients with post-COVID-19 sequelae (*25*). These studies offer preliminary contemporary evidence supporting the clinical use of formulated granules.

6.2. Japanese Kampo

Kampo medicine in Japan is deeply embedded within the national healthcare framework, with robust support from three pillars: widespread health insurance coverage, high rates of physician prescriptions, and a well-established system of real-world evidence.

Japan's health insurance system covers 148 ethical Kampo formulations and 187 crude drugs, with multiple clinical guidelines recommending Kampo as a treatment option (*26*). This has established an institutional foundation for its routine use. Surveys indicate that 92.7% of Japanese family physicians have prescribed Kampo formulations, frequently using them for patients with health-related anxiety or mental disorders (*27*). Of 679 clinicians, 30% regularly prescribe Kampo, while 45% prescribe it occasionally (*28*). A 2020 public survey revealed that 71% of respondents had taken Kampo medicines (*29*). Kampo is widely used for psychiatric disorders (*e.g.*, dementia, schizophrenia spectrum disorders, mood disorders, and anxiety disorders) (*30*) as well as cancer-related anorexia, general malaise, and peripheral neuropathy (*31*).

Common formulations and indications: Frequently prescribed medicines include Shakuyakukanzoto (an

analgesic, for neuralgia and joint pain) (*32*), Kakkonto (for acute upper respiratory tract inflammation) (*33*), Hochuekkito (for recovery from fatigue) (*34-36*), Rikkunshito and Daikenchuto (for upper and lower gastrointestinal discomfort) (*37,38*). Additionally, Yokukansan is frequently used to treat dementia and ICU delirium; Goreisan is used not only for gastrointestinal and renal diseases (*39,40*), but also for lymphedema following retroperitoneal lymph node resection (*41*); and formulations like Hachimijiogan have been found to enhance the quality of life for dialysis patients and potentially delay the initiation of dialysis, as evinced by studies on the impact of dialysis treatment on patient well-being and the importance of supportive care in improving outcomes (*42*). Data from a university hospital inpatient registry has indicated that Daikenchuto was prescribed most frequently (485 cases), primarily for postoperative and psychosomatic gastrointestinal symptoms (*43*).

According to the Yasui Classification, Kampo formulations are often combined with Western medications to enhance therapeutic efficacy (*e.g.*, Shichimitsukokato in combination with antihypertensive drugs), to mitigate adverse reactions to Western drugs (*e.g.*, Hochuekkito for treating severe diarrhea caused by azacytidine in patients with myelodysplastic syndromes), or to serve as substitutes when Western drugs are contraindicated (*e.g.*, Goreisan for treating headaches during pregnancy) (*44*).

The Evidence-Based Medicine Committee of the Japanese Society of Oriental Medicine has systematically screened over 500 randomized controlled trials (RCTs) and published the Kampo Best RCTs, providing a crucial clinical reference (*45*).

In summary, the Chinese model can be characterized as follows: personalized prescriptions serve as the foundation, a standardized system is in the process of development, and evidence-based data continues to accumulate. Its strengths reside in its strong clinical adaptability and its capacity to fully reflect the characteristics of TCM. Challenges include achieving unified regulation, ensuring consistent quality, and producing high-level evidence-based data. The Japanese model is characterized by a prerequisite of standardized prescriptions, a safeguard of unified medical insurance coverage, and a support system of evidence-based practice. Japanese Kampo medicine, rooted in TCM, has gained a significant market share globally, and particularly in Japan, due to its standardized production, consistent quality, and systematic clinical evidence. This approach has led to its popularity and dominance in the international market, with a focus on quality and safety. However, this standardized approach may limit the flexibility of tailoring prescriptions to individual symptoms.

7. Discussion

The different pathways to modernization in China and Japan as described in this review (formulated granules *vs.* Kampo medicines) are by no means merely a matter of differing technical choices; rather, they stem from divergences in regulatory philosophies, healthcare systems, and industrialization goals. An in-depth analysis of these institutional roots will help explain why the two models have diverged systematically in terms of quality standards, clinical evidence, and potential use internationally.

7.1. Structural contradictions between RCT design and consistent quality

The fixed formulation model of Japanese Kampo medicine essentially transforms traditional formulations into standardized industrial products. Its regulatory pathway requires that product ingredients, proportions, and manufacturing processes be completely fixed, making Kampo medicine naturally suited to classic placebo-controlled RCTs—where interventions are uniform and reproducible, in line with the paradigm of drug evaluation. However, this high level of reproducibility comes at the expense of the flexibility inherent in the core TCM principle of "differentiation of syndromes and treatment." In clinical practice, Japanese physicians rarely adjust formulations even when they observe changes in a patient's syndrome. While this ensures internal validity in RCTs, it undermines the need for individualized treatment in the real world.

In contrast, the "full-spectrum extraction plus clinical adjustment" model of formulated granules of TCM allows physicians to dynamically formulate, add, or subtract ingredients based on changing clinical presentations. While this preserves the flexibility of TCM, it poses a fundamental challenge to RCT design: how to create a comparable control group while maintaining individualized treatment? A deeper issue lies in the difficulty of ensuring the inter-manufacturer equivalence of formulated granules—differences in raw material sources, extraction processes, and excipients among manufacturers lead to significant variations in the chemical fingerprint profiles and dissolution profiles of the same standardized formulated granules. This not only diminishes the generalizability of multicenter RCT results but also prevents drug regulatory authorities from implementing uniformity controls as effectively as in Japan.

7.2. The cost of high reproducibility and clinical flexibility

The high reproducibility of Japanese Kampo medicines stems from its unique regulatory system of "pharmacy formulations combined with GMP uniformity." While this system ensures batch-to-batch consistency and the stability of Kampo medicines, it simultaneously imposes

institutional constraints on the prescribing freedom of Japanese Kampo physicians. While this model facilitates industrial production and evidence generation, it may lead to a dilution of therapeutic efficacy due to "inappropriate formulation-patient matching"—for example, the widespread use of Shosaikoto in Japan has been linked to cases of interstitial pneumonia due to insufficient syndrome differentiation (*48*). This illustrates that high reproducibility, when detached from clinical flexibility, may obscure the fundamental principles of traditional medicine.

Formulated granules of TCM offer immense clinical flexibility, but at the cost of an exponential increase in the complexity of source-level quality control. Since physicians are permitted to freely combine single-herb granules and an effectively infinite number of "potential formulations" exist on the market, but regulators can only control the quality of each type of granule once it is marketed and they cannot verify the final prescription as a whole. Due to differences in manufacturing processes, granules of the same type produced by different companies may exhibit varying bioequivalence, which directly leads to a crisis of prescription reproducibility across manufacturers and hospitals. This flexibility has become a fundamental obstacle to the international spread of formulated granules of TCM—particularly when entering markets such as the EU and the U.S., which require strict proof of bioequivalence for generic drugs.

There is essentially no perfect answer to how to balance the strengths and weaknesses of formulated granules of TCM and Japanese Kampo medicines. However, a potential breakthrough for Chinese formulated granules in the future may come from several approaches as described below.

Enhancing and unifying quality standards: Accelerating the development of a comprehensive system with quality standards and traceability that covers the entire process from raw materials to finished products. Focusing on resolving equivalence issues among products from different manufacturers, which arise due to variations in production processes, is a core task once national standards are fully implemented.

The systematic advancement of clinical evidence-based research can be informed by Japan's evidence-based path for development of Kampo medicine, and particularly by focusing on disease categories where Kampo has proven advantageous and on clinical formulations that are frequently utilized. More methodologically rigorous RCTs and real-world studies need to be designed and conducted to amass internationally recognized high-level medical evidence, providing robust support for the clinical use of TCM.

Holistic elucidation of compound mechanisms: Using integrative strategies such as multi-omics and bioinformatics, conducting more in-depth network pharmacology and systems pharmacology research into the synergistic mechanisms of granules with multiple components, targets, and pathways. Using modern scientific language to describe the scientific basis for their traditional efficacy.

Further exploring the potential for Sino-Japanese cooperation: Both countries are capable of achieving technological complementarity in areas including in-depth analysis of active ingredients, screening the material bases for broad-spectrum formulations, and collaboratively developing novel dosage forms. Additionally, they can jointly undertake targeted clinical research projects that comply with international standards.

Actively promoting international harmonization of regulations and standards: Jointly exploring strategies to evaluate traditional medicines within international frameworks like the World Health Organization, thereby removing obstacles for formulated granules of TCM to be registered internationally and to be allowed market access.

8. Conclusion

Formulated granules in TCM and Japanese Kampo preparations represent two distinct paradigms in the modernization of traditional medicine. Formulated granules, rooted in the traditional theory of syndrome differentiation and treatment, serve the vast and complex clinical needs of the Chinese population. Formulated granules are developing on a path toward industrial production and flexible clinical formulations, and a standardized system of those formulations is being created. Japanese Kampo, which relies on fixed prescriptions and a universal healthcare system, has enabled highly standardized and intensive production as well as evidence-based use. Although the paths of their development diverge, both face the common challenge of demonstrating efficacy and safety within a modern scientific framework while ensuring precise, end-to-end quality control. The future development of formulated granules of TCM relies on maintaining the core principles of TCM theory while also enhancing quality standards, advancing pharmacodynamic research, and amassing persuasive clinical evidence. Rationally learning from Japan's mature experience in Kampo medicine and particularly standardization, evidence-based medical practice, and the close integration of industry and healthcare, TCM can also draw on the unique advantages of formulated granules in terms of clinical flexibility and technological innovation. TCM can offer significant practical and long-term strategic value by contributing to global health.

Acknowledgements

The author wishes to thank professor Wei Tang for his insightful discussion.

Funding: This work was supported by a grant from

the Qinghai Province Kunlun Elite Innovation and Entrepreneurship Project.

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Wang WY, Xie Y, Zhou H, Liu L. Contribution of traditional Chinese medicine to the treatment of COVID-19. Phytomedicine. 2021; 85:153279.
2. NobelPrize.org. The Nobel Prize in Physiology or Medicine 2015. Nobel Media AB, 2015. https://www.nobelprize.org/prizes/medicine/2015/summary/.
3. Wang WY, Zhou H, Wang YF, Sang BS, Liu L. Current policies and measures on the development of traditional Chinese medicine in China. Pharmacol Res. 2021; 163:105187.
4. Zhang C, Zhang CG, Liu M, Wang T. Analysis of the impact of extraction processes on the quality of traditional Chinese medicine granules. Phytochem. Anal. 2025; 36:903-911.
5. Lu L, Shi JH, Hou FG, Wang XP, Xie MD, Yue YS, Gui XJ, Li XL, Liu RX. Chinese medicine dispensing granules: History, status quo, and development prospect in "post-pilot era". Zhongguo Zhong Yao Za Zhi. 2022; 47:2008-2014. (in Chinese)
6. Gao RL, Shi GH, Fu J, Jia YH, Zhang ZQ. History of development of and future prospects for formulated granules of traditional Chinese medicines. Zhongguo Chu Fang Yao. 2026; 24:117-121. (in Chinese)
7. Motoo Y, Seki T, Tsutani K. Traditional Japanese medicine, Kampo: Its history and current status. Chin J Integr Med. 2011; 17: 85-87.
8. Hakamatsuka T. Marketing Approval of ethical kampo medicines. Yakugaku Zasshi. 2017; 137:163-165.
9. Committee for Research and Development of Educational Programmes in Medicine. Model Core Curriculum for Medical Education. Tokyo: Ministry of Education, Culture, Sports, Science, and Technology, 2001; pp. 1-61.
10. Chen SL, Liu CX, Zhang TJ, Liu A, Zhu GW, Wang YW, Chen CQ, Zhang HB, Xiao XF, Huang YH. Ideas and suggestions on CMM decoction inheritance based on CMM quality markers and traditional usage. Zhong Cao Yao. 2019; 50:4519-4528. (in Chinese)
11. Cui T, Hou Y, Feng H, Wu S, Li W, Li Z. Granulation process analysis technologies and potential applications in traditional Chinese medicine. Acupunct Herb Med. 2022; 2:9-24.
12. Maegawa H, Nakamura T, Saito K. Regulation of traditional herbal medicinal products in Japan. J Ethnopharmacol. 2014; 158:511-515.
13. Hagihara K, Abbreviation of Kampo Formulations G, Yakubo S, Committee for Vocabulary in the Japan Society for Oriental M, Namiki T. Abbreviation of kampo formulations and basic terminology in kampo medicine. Trad Kampo Med. 2017; 4:65-88.
14. Lai E, Ning J, Liu Y, Zhang Y, Luo W, Hu Y, Liu Y, Liang Z, Li P, Tian J, Zheng G, Wei M. The chemical profile and the quantification of the traditional Chinese medicine formula granules Erchen Decoction analyzed by ultra-high performance liquid chromatography-Q-exactive orbitrap tandem mass spectrometry method. Sep Sci Plus. 2024; 7:e202400142.
15. Wang R, Yang T, Feng Q, *et al*. Integration of network pharmacology and proteomics to elucidate the mechanism and targets of traditional Chinese medicine Biyuan Tongqiao granule against allergic rhinitis in an ovalbumin-induced mice model. J Ethnopharmacol. 2024; 318:116816.
16. Sun X, Xiang H, Liu Z, Xiao H, Li X, Gong W, Pan L, Zhao L, Yao J, Sun C, Zhang G. Jingfang Granules alleviates bleomycin-induced acute lung injury through regulating PI3K/Akt/mTOR signaling pathway. J Ethnopharmacol. 2024; 318:116946.
17. Zhang J, Li W, Xue S, Gao P, Wang H, Chen H, Hong Y, Sun Q, Lu L, Wang Y, Wang Q. Qishen granule attenuates doxorubicin-induced cardiotoxicity by protecting mitochondrial function and reducing oxidative stress through regulation of Sirtuin3. J Ethnopharmacol. 2024; 319:117134.
18. Wu Z, Zhang Y, Mao D, Wang N, Lu Z, Yan H, Lan Y, Wang M, Zhang R, Peng M, Zeng G. Integrated proteomics and metabolomics analysis of liver tissue reveal the therapeutic effect of Jiedu Huayu Granules on acute liver failure in rats. J Pharm Biomed Anal. 2025; 266:117072.
19. Chen J, Chen Q, Xiao P, Jin W, Yu L. A novel framework for uncovering the coordinative spectrum-effect correlation of the effective components of Yangyin Tongnao Granules on cerebral ischemia-reperfusion injury in rats. J Ethnopharmacol. 2025; 337(Pt 1):118844.
20. Tokita Y, Sugaya K, Nishijima S, Tsuchiya N, Hamaguchi M, Yamamoto H. Choreito, a Kampo medicine attenuates detrusor overactivity and bladder pain symptoms in rat tranilast-induced interstitial cystitis/bladder pain syndrome-like model. Neurourol Urodyn. 2023; 42:56-64.
21. Takei M, Tokita Y, Hamaguchi M, Aishima M, Ichikura S, Okabe A, Naito S, Yokomizo A. A Kampo medicine, choreito, evaluated on female lower urinary tract symptoms (LUTS) with bladder pain. Trad Kampo Med. 2025; 12:47-54.
22. Mori C, Yamamoto R, Shirafuji N, Asano R, Sasaki H, Yamaguchi T, Kitazaki Y, Endo Y, Enomoto S, Hamano T. Traditional Japanese Kampo Medicine Yokukansan decreases phosphorylated tau and oligomeric tau. Alzheimers Dement. 2024; 20:e093352.
23. Sato-Numata K, Suzuki T, Saito H, Kato S, Sakai A, Mori S, Nakae H, Hasegawa H, Okada Y, Numata T. Boi-Ogi-To, a traditional Japanese kampo medicine, promotes cellular excretion of chloride and water by activating volume-sensitive outwardly rectifying anion channels. FASEB J. 2025; 39:e70573.
24. Chen B, Yu X, Zhang L, Huang W, Lyu H, Xu Y, Shen J, Yuan W, Fang M, Li M, Gao Y. Clinical efficacy of Jingyin granules, a Chinese patent medicine, in treating patients infected with coronavirus disease 2019. Phytomed. 2023; 108:154496.
25. Pang W, Yang F, Zhao Y, *et al*. Qingjin Yiqi granules for post-COVID-19 condition: A randomized clinical trial. J Evid Based Med. 2022; 15:30-38.
26. Uneda K, Yoshino T, Ito H, Imoto S, Nogami T. Current situation and future issues with Kampo medicine: A survey of Japanese hysicians. Trad Kampo Med. 2024; 11:156-166.
27. Takayama S, Akaishi T, Nozaki H. Characteristics and course of patients treated with Kampo medicine in the Department of General Medicine. J Gen Fam Med. 2020; 21:48-55.

28. Muramatsu S, Aihara M, Shimizu I, Arai A, Kajii E. Current status of kampo medicine in community health care. Gen Med. 2012; 13:37-45.
29. Chuang CY, Chung HY, Uehara C, Taira M, Takeda O, Nagashima Y, Honda M, Chen FS, Arai I. A survey of Japanese people's attitude to and experience of Kampo medicines. Trad Kampo Med. 2023; 10:35-43.
30. Tatsumi L, Suzuki T, Yamada K, Mimura M, Uchida H. Kampo, A Japanese traditional medicinal system for psychiatric conditions: A narrative review. Pharmacopsych. 2019; 52:251-260.
31. Motoo Y, Cameron S. Significance of Kampo medicine for cancer supportive care: Overview. Trad Kampo Med. 2023; 10:16-19.
32. Kubota T, Miyata A. Successful use of Shakuyaku-kanzo-to, a traditional herbal medicine, for intractable symptoms of thoracic outlet syndrome: A case report. J Anesth. 2005; 19:157-159.
33. Kurokawa M, Tsurita M, Brown J, Fukuda Y, Shiraki K. Effect of interleukin-12 level augmented by Kakkon-to, a herbal medicine, on the early stage of influenza infection in mice. Antiviral Res. 2002; 56:183-188.
34. Kimura M, Sasada T, Kanai M, Kawai Y, Yoshida Y, Hayashi E, Iwata S, Takabayashi A. Preventive effect of a traditional herbal medicine, Hochu-ekki-to, on immunosuppression induced by surgical stress. Surg Today. 2008; 38:316-322.
35. Tsuge A, Chiba S, Yagura Y, Okamoto M, Muto S, Hisaka S, Nose M. Hochuekkito exerts the anti-allergic effects *via* activating regulatory T cells in a murine model of contact hypersensitivity. J Nat Med. 2023; 77:352-362.
36. Hamada H, Sekikawa K, Murakami I, Aimoto K, Kagawa K, Sumigawa T, Okusaki K, Dodo T, Awaya Y, Watanabe M, Kondo K, Ogawa T, Yamamoto H, Hattori N. Effects of Hochuekkito combined with pulmonary rehabilitation in patients with chronic obstructive pulmonary disease. Exp Ther Med. 2018; 16:5236-5242.
37. Oka T, Tamagawa Y, Hayashida S, Kaneda Y, Kodama N, Tsuji S. Rikkunshi-to attenuates adverse gastrointestinal symptoms induced by fluvoxamine. Biopsychosoc Med. 2007; 1:21.
38. Manabe N, Camilleri M, Rao A, Wong BS, Burton D, Busciglio I, Zinsmeister AR, Haruma K. Effect of daikenchuto (TU-100) on gastrointestinal and colonic transit in humans. Am J Physiol Gastrointest Liver Physiol. 2010; 298:G970-975.
39. Morita F, Yokokawa H, Matsuda N, Fujibayashi K, Uehara Y, Kobayashi H, Naito T. Comparative efficacy of goreisan and probiotics in Japanese adults with acute infectious gastroenteritis: Randomized controlled trial. Trad Kampo Med. 2017; 4:89-93.
40. Sugihara R, Hashimura T, Sakata Y. Goreisan as a successful adjuvant therapy of heart failure with preserved ejection fraction and advanced chronic kidney disease: A case report. Oxf Med Case Reports. 2023; 10:omad116.
41. Komiyama S, Takeya C, Takahashi R, Yamamoto Y, Kubushiro K. Feasibility study on the effectiveness of Goreisan-based Kampo therapy for lower abdominal lymphedema after retroperitoneal lymphadenectomy *via* extraperitoneal approach. J Obstet Gynaecol Res. 2015; 41:1449-1456.
42. Shimizu M, Takayama S, Ishizawa K, Abe M, Ishii T. Kampo medicine can improve quality of life and prolong hemodialysis implementation in patients with advanced-stage chronic kidney disease. Trad Kampo Med. 2021; 8:229-233.
43. Sugimine R, Kikukawa Y, Kurihara D, Arita R, Takayama S, Kikuchi A, Ohsawa M, Ishii T. Kampo medicine prescriptions for hospitalized patients in Tohoku University Hospital. Trad Kampo Med. 2021; 8:221-228.
44. Motoo Y, Yasui H. Analysis of Kampo case reports from the viewpoint of "Yasui Classification." Trad Kampo Med. 2024; 11:60-64.
45. Motoo Y, Tominaga K, Suzuki H, Kaido T, Matsuda C, Morita S, Kogure T, Arai I, Tsutani K. Selection of best randomized controlled trials of Japanese Kampo medicines: Introduction of the trials and the hardships in their conduction. Trad Kampo Med. 2025; 12:63-73.
46. Chinese Pharmacopoeia Commission. The Pharmacopoeia of the Peoples Republic of China. China Medical Science and Technology Press. Beijing, China, 2025.
47. Ministry of Health, Labor, and Welfare. General Recognition Criteria for Kampo Preparations. Tokyo, Japan. 2012.
48. Sato A, Toyoshima M, Kondo A, Ohta K, Sato H, Ohsumi A. Pneumonitis induced by the herbal medicine Sho-saiko-to in Japan. Nihon Kyobu Shikkan Gakkai Zasshi. 1997; 35:391-395. (in Japanese)

Received February 28, 2026; Revised April 16, 2026; Accepted April 18, 2026.

**Address correspondence to:*
Wenyuan Li, School of Pharmacy, Qinghai University, 251 Ningda Road, Chengbei District, Xining, China 810016.
E-mail: qhliwenyuan@126.com

Released online in J-STAGE as advance publication April 24, 2026.

Original Article

DOI: 10.5582/ddt.2026.01009

Amelioration of gastrointestinal motility and gut dysbiosis by acupoint application of Tongfu Powder to loperamide-induced constipation in mice

Ke Wang[1,2], Pingping Cai[1], Lin Zhao[1], Zhixue Wang[1], Yuqi Wang[1], Fanghua Qi[1,*]

[1] Traditional Chinese Medicine, Shandong Provincial Hospital affiliated to Shandong First Medical University, Ji'nan, China;
[2] Xuzhou Tongshan District Hospital of Traditional Chinese Medicine, Xu'zhou, China.

SUMMARY: Chronic constipation is one of the most common gastrointestinal disorders worldwide. It has an impact on daily life and poses a considerable economic burden. Tongfu Powder originated from the Xiaochengqi Decoction with the intent of promoting defecation. Acupoint application of Tongfu Powder has exhibited potentially beneficial effects in the treatment of constipation. However, the potential mechanisms by which acupoint application of Tongfu Powder regulates gastrointestinal motility and gut microbiota are still unclear. The current study sought to investigate the effects and underlying mechanisms of acupoint application of Tongfu Powder on loperamide-induced constipation in mice. The results demonstrated that acupoint application of Tongfu Powder significantly improved the overall defecation of constipated mice, including an increase in the number of fecal pellets, fecal weight, and water content, a decrease in gastric residual volume, and an increase in the intestinal propulsion rate. It also alleviated loperamide-induced colonic histopathological deterioration such as cellular infiltration and thinning of the muscular and mucosal layers in constipated mice. Acupoint application of Tongfu Powder significantly up-regulated the levels of interstitial cells of Cajal (ICC) markers (c-Kit and SCF) and it increased synthesis of intestinal 5-hydroxytryptamine (5-HT) and related proteins (TPH1, HTR4 and SERT). Acupoint application of Tongfu Powder promoted intestinal mucin-2 (MUC2) secretion and increased the expression of tight junction proteins (claudin-1 and occludin). 16S rRNA gene sequencing revealed that acupoint application of Tongfu Powder significantly increased the abundance of *Akkermansia muciniphila*, a bacterium known to be involved in regulating gut motility and intestinal barrier function, thereby alleviating intestinal dysfunction. In addition, it is worth noting that the therapeutic effect of Tongfu Powder acupoint application combined with lactulose is superior to that of either treatment alone. In conclusion, results revealed that acupoint application of Tongfu Powder might alleviate loperamide-induced constipation by regulating the intestinal barrier and gut microbiota.

Keywords: Chronic constipation, Tongfu powder, acupoint application, gastrointestinal motility, gut microbiota

1. Introduction

Chronic constipation is one of the most frequent gastrointestinal disorders caused by altered gastrointestinal dynamics. Its incidence is increasing due to modern lifestyles, dietary composition, and psychological changes, with a global prevalence between 10% and 15% in the population (*1*). Chronic constipation is not a life-threatening condition, but it has a great impact on the quality of life of patients, posing a serious mental and physical burden (*2*). In addition, chronic constipation can increase economic costs and result in a burden to healthcare delivery systems worldwide (*3*). Direct costs attributed to constipation-related health care in the US are estimated to be more than US$230 million per year (*4*).

Various treatment options are generally recommended by the American Gastroenterological Association, including dietary changes, lifestyle modifications, the use of fiber supplements, stool softeners, and laxatives (*5*). Unfortunately, more than 50% of patients fail to respond to these standard treatments mainly due to dissatisfaction with efficacy, safety, adverse reactions, and cost (*6*). Therefore, the treatment of chronic constipation remains a challenge. A safe and cost-effective treatment for chronic constipation needs to be identified. Nowadays, many patients with chronic constipation seek help from complementary and alternative therapies, and external treatment with traditional Chinese medicine (TCM) is their usual choice (*7*).

External treatment with TCM, as a unique traditional treatment with a long history in China, refers to a series

of external treatments such as acupuncture, moxibustion, massage, and acupoint application (*8*). These external therapies were reported to be efficacious and to avoid adverse consequences such as abdominal pain, electrolyte disturbance, melanosis coli, and severe drug dependence after long-term use of laxatives (*9*). Acupoint application is a form of external TCM that is considered a non-invasive method combining acupuncture points, meridians, and Chinese herbal medicines. Chinese herbal medicines can treat or prevent diseases by permeating and stimulating skin-related acupoints (*10*). When treating constipation, several Chinese herbal medicines to promote defecation are usually the preferred choice for acupoint application.

Xiaochengqi Decoction is an ancient traditional herbal formula originating from the Treatise on Febrile Diseases (Han Dynasty China), which has been used to treat chronic constipation for thousands of years (*9*). It consists of three herbs, *Rheum officinale* Baill., *Citrus aurantium* L., and *Magnolia officinalis* Cortex. Tongfu Powder is a prescription formulated by the authors' hospital based on the Xiaochengqi Decoction and consists of five herbs, *Rheum officinale* Baill., *Citrus aurantium* L., *Magnolia officinalis* Cortex, *Areca catechu* L. and *Dryobalanops aromatica* C.F. Gaertn. The plant names have been checked with "World Flora Online (WFO)" (*http://www.worldfloraonline.org*). *Rheum officinale* Baill. extract can promote the secretion of colonic mucus and regulate the intestinal flora related to mucin secretion (*11*), while *Citrus aurantium* L. can increase colon propulsion and correct abnormal colonic slow waves to relieve chronic constipation (*12*). *Magnolia officinalis* Cortex can significantly promote gastrointestinal motility, and its main component magnolol can alleviate gastrointestinal dysfunction by up-regulating the expression of the c-kit/SCF signal pathway (*13*). Arecoline, as the main active component of *Areca catechu* L., can stimulate the contraction of rat distal colon smooth muscle *via* muscarinic (M3) receptor-mediated extracellular calcium influx and calcium pool release (*14*). It also can relieve constipation by regulating intestinal microorganisms and their metabolites and by regulating the intestinal genome (*15*). In addition, *Dryobalanops aromatica* C.F. Gaertn., as a natural transdermal enhancer, can effectively promote the absorption of herbal active ingredients through the skin (*16*).

Tongfu Powder has significant laxative effects and has been used for acupoint application to treat chronic constipation in clinics for more than ten years at the authors' hospital, but the underlying mechanisms by which acupoint application of Tongfu Powder treats chronic constipation remain unknown. A point worth noting is that gut microbiota dysbiosis is closely associated with the pathogenesis of gut dysmotility, which is receiving increasing attention (*17*). Gut microbiota dysbiosis results in inflammatory responses, an impairment in colonic epithelial integrity, and a reduction in bowel secretions, thus decreasing intestinal peristalsis and inducing constipation (*18*). Therefore, the current study aimed to explore the potential mechanisms of acupoint application of Tongfu Powder in treating chronic constipation in the context of the intestinal barrier and microbiota. The hope is that this experimental study will provide new avenues and theoretical support for the clinical practice of acupoint application of Tongfu Powder in treating chronic constipation.

2. Materials and Methods

2.1. Preparation of Chinese herbal medicines and acupoint selection for acupoint application

The Chinese herbal medicine Tongfu Powder used in this study originated from the Xiaochengqi Decoction, and its preparation was as follows: *Rheum officinale* Baill., *Citrus aurantium* L., *Magnolia officinalis* Cortex, *Areca catechu* L. and *Dryobalanops aromatica* C.F. Gaertn were ground into powder, mixed in a ratio of 3:5:5:5:1, and then made into a paste with Chinese yellow rice wine. The above drugs are all from the traditional Chinese medicine pharmacy of Shandong Provincial Hospital. Approximately 1 g of the Chinese herbal medicine ointment was placed onto a ring-shaped patch (with a diameter and thickness of about 1.5 cm and 0.5 cm, respectively) and applied to the navel center (Shenque point) of mice (Figure 1A). Shenque (RN8), located at the navel, is one of the effective acupuncture points for chronic constipation (*19*).

2.2. Chemicals

A lactulose oral solution was purchased from Abbott Biologicals B.V. Loperamide hydrochloride (L129465-1g) was purchased from Aladdin (Shanghai, China). A serotonin 5-hydroxytrytamine (5-HT) ELISA kit was purchased from KeyBio (Jinan, China). The mouse monoclonal antibody SCF (sc-13126) was purchased from Santa Cruz Biotechnology, Inc. (Dallas, Texas). The rabbit monoclonal antibodies C-Kit (ab256345) and TPH (ab52954) were purchased from Abcam (Cambridge, UK). Claudin-1 (CY6872) and Occludin (CY5997) were purchased from Abways (Shanghai, China). SERT (19559-1-AP), HTR-4 (21165-1-AP), and MUC2 (27675-1-AP) were purchased from PTG (Wuhan, China). GAPDH (ZB15004-HRP-100) was purchased from Servicebio (Wuhan, China).

2.3. Animals and experimental design

Forty specific-pathogen-free (SPF) male Kunming mice, aged 8 weeks and weighing 40–47 g, were purchased from Beijing Charles river Laboratory Animal Technology Co., Ltd. (Beijing, China; license No.

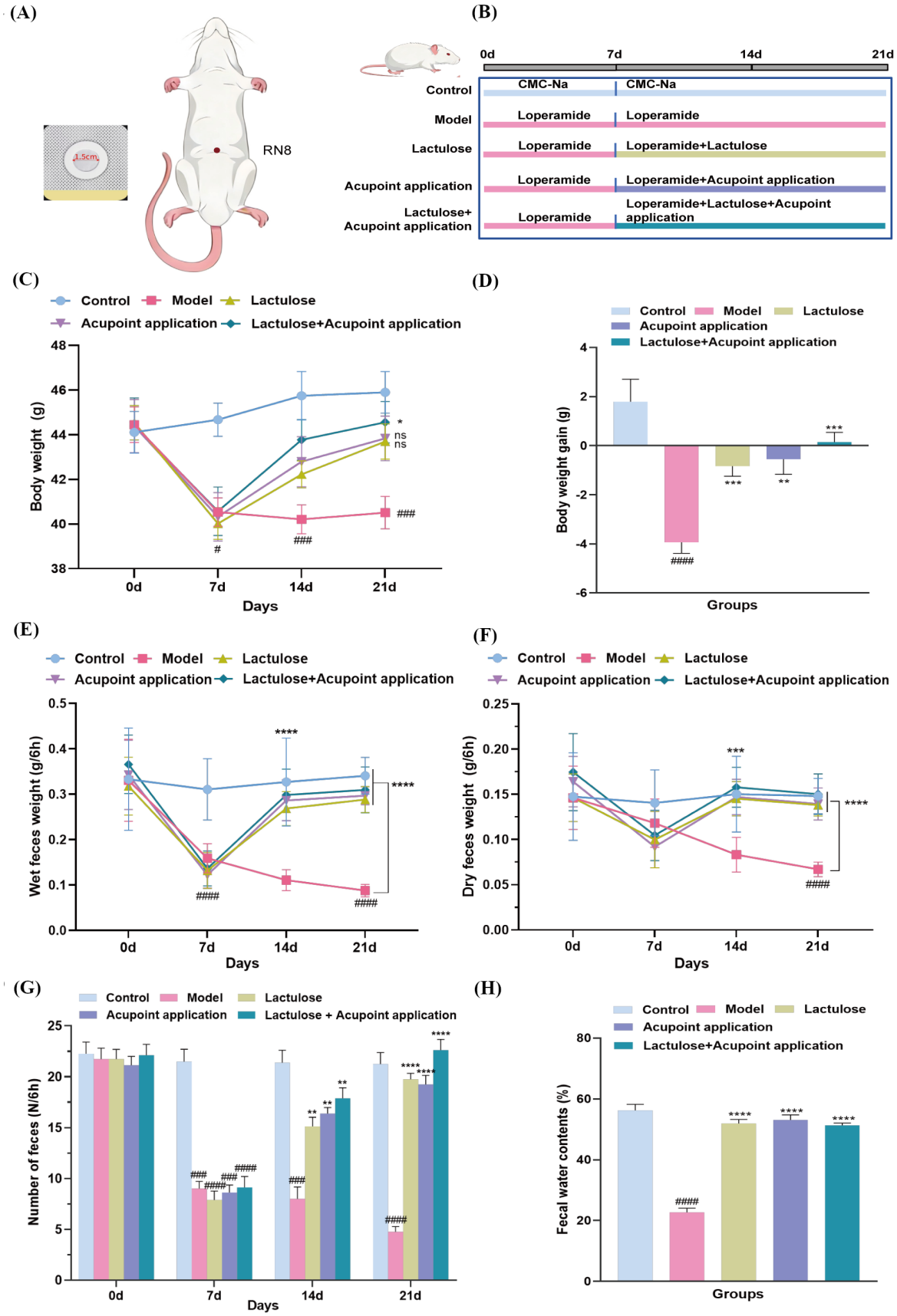

Figure 1. Study setup and effects of acupoint application on physiological parameters in constipated mice. (**A**) Tool for acupoint application and map of the abdominal Shenque acupoint (RN8) in mice. (**B**) Schematic diagram of the experimental design. (**C**) Body weight. (**D**) Body weight gain on the 21st day of the experiment compared to that on the 1st day. (**E-H**) Changes in the fecal wet weight, dry weight, pellet number, and water content within 6 hours. Data were expressed as the mean ± SEM ($n = 8$). $^{\#}p < 0.05$, $^{\#\#\#}p < 0.001$, and $^{\#\#\#\#}p < 0.0001$ *vs.* the control group; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, and $^{****}p < 0.0001$ *vs.* the model group; ns, not significant.

SCXK(Jing)2021-0011). The mice were housed in an environment with a room temperature ranging from 20 to 24°C, a humidity level of 50 to 70%, and a 12-hour light/dark cycle. They were provided with free access to standard laboratory food and water. All procedures followed the management guidelines set forth by the Experimental Animal Ethics Committee of Shandong Provincial Hospital affiliated with Shandong First Medical University (Ethics No. 2022-096). In addition, all methods used in this study are in accordance with the ARRIVE Guidelines.

After one week of adaptive feeding, the mice were randomly divided into five groups based on their body weight with 8 mice in each group: the control group (Control), the constipation model control group (Model), the lactulose group (Lactulose), the acupoint application of Tongfu Powder group (Acupoint application), and the group treated with a combination of lactulose and acupoint application (Lactulose+Acupoint application).

A mouse model of constipation was established by slightly modifying the method used in a previous study (*20*). Because loperamide is insoluble in water, it is suspended in 0.5% sodium carboxymethyl cellulose (CMC-Na) for subsequent administration by gavage. During the experiment, from day 1 to day 21, all mice (except the control group) were given 0.2 mL of loperamide (10 mg/kg) by gavage at 8 o'clock every day to induce and maintain constipation. Mice in the control group were given the same dose of 0.5% CMC-Na. After the constipation model was successfully established, from the 8th day to the 21st day, at 10 o 'clock every day, the mice in the three treatment groups received the following treatment: the lactulose group was given 0.2 mL of a lactulose solution (3.9 mL/kg), the acupoint application group was treated with herbal medicines for 6 hours, and the combined treatment group was treated with both lactulose and acupoint application. The timeline for each treatment group is depicted in Figure 1B.

2.4. Determination of body weight and defecation

On the 0th and 7th day, the weight of each mouse, the number and weight of fecal pellets within 6 hours, and the moisture content of feces were recorded. To calculate the moisture content, the fresh feces of mice were dried in a ventilated oven at 60°C until a constant weight was obtained. Then, the following formula was used to determine the fecal moisture content: Fecal moisture content (%) = (difference between wet weight and dry weight) / wet weight × 100%. These indicators were used to evaluate the defecation of mice and determine whether the constipation model was successfully established. On days 14 and 21, the body weight and feces of all mice were recorded again to evaluate the defecation of mice.

In order to evaluate the gastrointestinal transport capacity of mice, 0.4 mL of a black semi-solid paste (consisting of water, CMC-Na, milk powder, sugar, starch, and an ink mixture) was administered intragastrically after a 12-hour fasting period following the last administration. After 20 minutes of intragastric administration, the mice were immediately sacrificed. The gastrointestinal transit rate was calculated by determining the percentage of the distance the ink in the semi-solid paste moved in relation to the length of the entire small intestine. In addition, the stomach of mice was weighed to evaluate the gastric emptying of mice.

2.5. Tissue staining

After the mice were sacrificed, the colon tissue was gently shaken in pre-cooled PBS to remove the intestinal contents. A portion of the colon was preserved in liquid nitrogen for subsequent experiments, while the other part was fixed in a 4% paraformaldehyde fixed solution. The colon tissue of mice was fixed for 48 hours, embedded in paraffin, and cut into 4-μm sections for section staining and immunohistochemistry. Hematoxylin-eosin staining (H&E) was used to evaluate histopathological changes, and Alcian blue (AB) staining was used to evaluate the amount of colonic mucus. Finally, the histopathological morphology was observed with an optical microscope and quantified using the software Image J 1.54d.

2.6. Immunohistochemical (IHC) staining

To analyze the expression of MUC2 protein, IHC staining was used as mentioned in our previous studies (*21*). The colon tissue slices were dewaxed and hydrated and then soaked in 0.3% methanol hydrogen peroxide for 30 minutes to block endogenous peroxidase activity. Next, the slices were incubated overnight at 4°C with primary antibody MUC2 (1:1,000), followed by 1 hour of incubation at room temperature with the anti-rabbit secondary antibody. Staining was visualized using the DAB horseradish peroxidase chromogenic kit (Wuhan Service Biotechnology Co., Wuhan, China), and samples were counterstained with hematoxylin. The expression of MUC2 was detected and photographed using the 3D HISTECH Digital Pathology System. The average optical density (MOD) was measured using the graphic analysis software Image J. All IHC staining was independently evaluated by two researchers, and any discrepancies were resolved through group discussion.

2.7. Enzyme-linked immunosorbent assay (ELISA)

To assess the level of the neurotransmitter 5-HT, blood was collected from the eyeballs of anesthetized mice using a sterile tube. The blood was then left at room temperature for 2 hours before being centrifuged at 3,500 rpm at 4°C for 15 minutes to obtain serum samples. All serum samples were packaged and stored in a −80°C refrigerator. The level of 5-HT in serum was determined using an ELISA kit, as mentioned above (*22*).

Table 1. Sequences of primers used for amplification and sequencing

Primer	Sequence (5'-3')	
	Forward	Reverse
C-kit	GACCCGACGCAACTTCCTTAT	TGGCAGCATCCGACTTAATCAA
SCF	TCTGCGGGAATCCTGTGACT	CGGCGACATAGTTGAGGGTTAT
Claudin-1	AGCACCGGGCAGATACAGT	GCCAATTACCATCAAGGCTCG
Occludin	CTCCTCCAATGGCAAAGTGAATG	GTCATCCACACTCAAGGTCAGA
Gapdh	TGCACCACCAACTGCTTAG	GATGCAGGGATGATGTTC

2.8. Real-time quantitative PCR (RT-qPCR)

The expression of mRNA of several genes was detected with RT-qPCR as described previously (*23*). Total RNA was extracted from mouse colon tissue using the TRIzol reagent (Exact Biology, Hunan, China), and 1 μg of RNA was reverse-transcribed into complementary DNA (cDNA) according to the instructions for AbScript III RT Master Mix (ABclonal Technology, Wuhan, China). Then, SYBR Green (ABclonal Technology, Wuhan, China) was used as a fluorescent marker, and 2 μL cDNA and corresponding upstream and downstream primers (10 μmol/L) were added to the 20 μL reaction system and amplified using the Fluorescence Quantitative PCR Detection System (Bioer, Hangzhou, China). The sequences of amplification primers (GAPDH, SCF, c-kit, claudin-1, and occludin) are shown in Table 1. The amplification protocol was as follows: initiation at 37°C for 2 m, 95°C for 3 m, one cycle each, and then amplification at 95°C for 40 cycles for 5 s and 60°C for 30 s. GAPDH was used as a reference gene, and the relative expression of the target gene was calculated according to the $2^{-\Delta\Delta Ct}$ method.

2.9. Western blotting (WB)

As mentioned previously (*24*), the expression of several proteins was detected with WB. After extracting total protein from colon tissue, proteins with different molecular weights were separated using gel electrophoresis. After the WB membrane was sealed in QuickBlock™ for 15 minutes, it was incubated with primary antibodies at 4°C overnight, including c-kit (1:1,000), SCF (1:2,000), HTR4 (1:1,000), Tph1 (1:1,000), and SERT (1:1,000). Then, the membrane was incubated with the secondary antibody for 1 h at room temperature. Finally, the protein bands were detected using enhanced chemiluminescence.

2.10. 16S rDNA sequencing of intestinal bacteria

DNA was extracted from the intestinal contents of mice using a CTAB kit. The V3-V4 variable region was amplified using barcode primers (341F (5'-CCTACGGGGNGGCWGCAG-3') and 805R (5'-GACTACHVGGGGTATCTAATCC-3')). PCR products were purified using AMPure XT beads (Beckman Coulter Genomics, Danvers, MA, USA) and quantified using Qubit (*Invitrogen*, USA). The purified PCR products were evaluated using the Agilent 2100 Bioanalyzer (Agilent, USA) and Illumina (Kapa Biosciences, Woburn, MA, USA) library quantification kits. The qualified library concentration should be above 2 nM. After gradient dilution of the qualified sequencing libraries (Index sequence is not repeatable), they were mixed in the corresponding proportion according to the required sequencing amount and denatured with NaOH to a single strand for sequencing. The NovaSeq 6000 sequencing instrument was used for 2 × 250 bp double-ended sequencing, and the corresponding reagent was NovaSeq 6000 SP Reagent Kit (500 cycles). After sequencing, all data were analyzed.

2.11. Statistical analyses

The software GraphPad Prism 8.0 (GraphPad Software, LA Jolla, CA, USA) was used to perform statistical analysis, and the results were expressed as the mean ± SEM. One-way analysis of variance (ANOVA) was used to determine the differences between more than two groups, and comparisons between two groups were further analyzed using Tukey's multiple comparisons test. For all results, statistical significance was indicated by a p value < 0.05.

3. Results

3.1. Effects of acupoint application on improving defecation in constipated mice

Throughout the experiment, the mice exhibited a sluggish response, reduced activity, yellow and dull fur, anorexia, and other symptoms after administration of loperamide. Their feces were dry, short, hard, and light yellow, with undigested food visible in them. However, mice in the control group were active and had white, shiny fur and wet, sticky feces that were slightly larger, softer, and darker.

As shown in Figure 1C-1D, on the 7th day of loperamide treatment, loperamide significantly inhibited the weight gain of mice compared to the control group, and the number of fecal pellets, dry and wet weight,

and fecal water content of mice decreased significantly (Figure 1E-1H). This indicates that the constipation model has been successfully replicated. After two weeks of treatment, the weight of mice in the lactulose group, acupoint application group, and lactulose + acupoint application group increased, and defecation gradually recovered compared to that in the model group.

In addition, the small intestine propulsion rate in the three intervention groups improved significantly compared to that in the model group, and the gastric residual weight decreased significantly (Figures 2A-2C). Moreover, there were no significant differences in the intestinal propulsion rate and gastric residual weight between the acupoint application group and the lactulose group ($p > 0.05$). Compared to the lactulose group or the acupoint application group, the lactulose + acupoint application group had a significantly higher intestinal propulsion rate ($p < 0.05$). These results show that the acupoint application of Tongfu Powder alleviated the symptoms of constipation in mice and maintained the healthy state of the gastrointestinal tract by increasing the amount and moisture content of feces and accelerating gastrointestinal peristalsis. The efficacy of the lactulose group was similar to that of the acupoint application group, whereas the combination therapy exhibited superior effects.

3.2. Effects of acupoint application on reducing colon tissue damage

To evaluate the effects of acupoint application of Tongfu Powder on the intestinal composition of constipated mice, H&E staining was used. Compared to the control group, staining in the model group indicated that loperamide caused a decrease in the number and disordered arrangement of goblet cells in the colon of mice, edema in the submucosa, destruction of the crypt structure, and thinning of muscle and mucosal layers (Figure 3A). Fortunately, the damaged colonic structures in the lactulose group, acupoint application

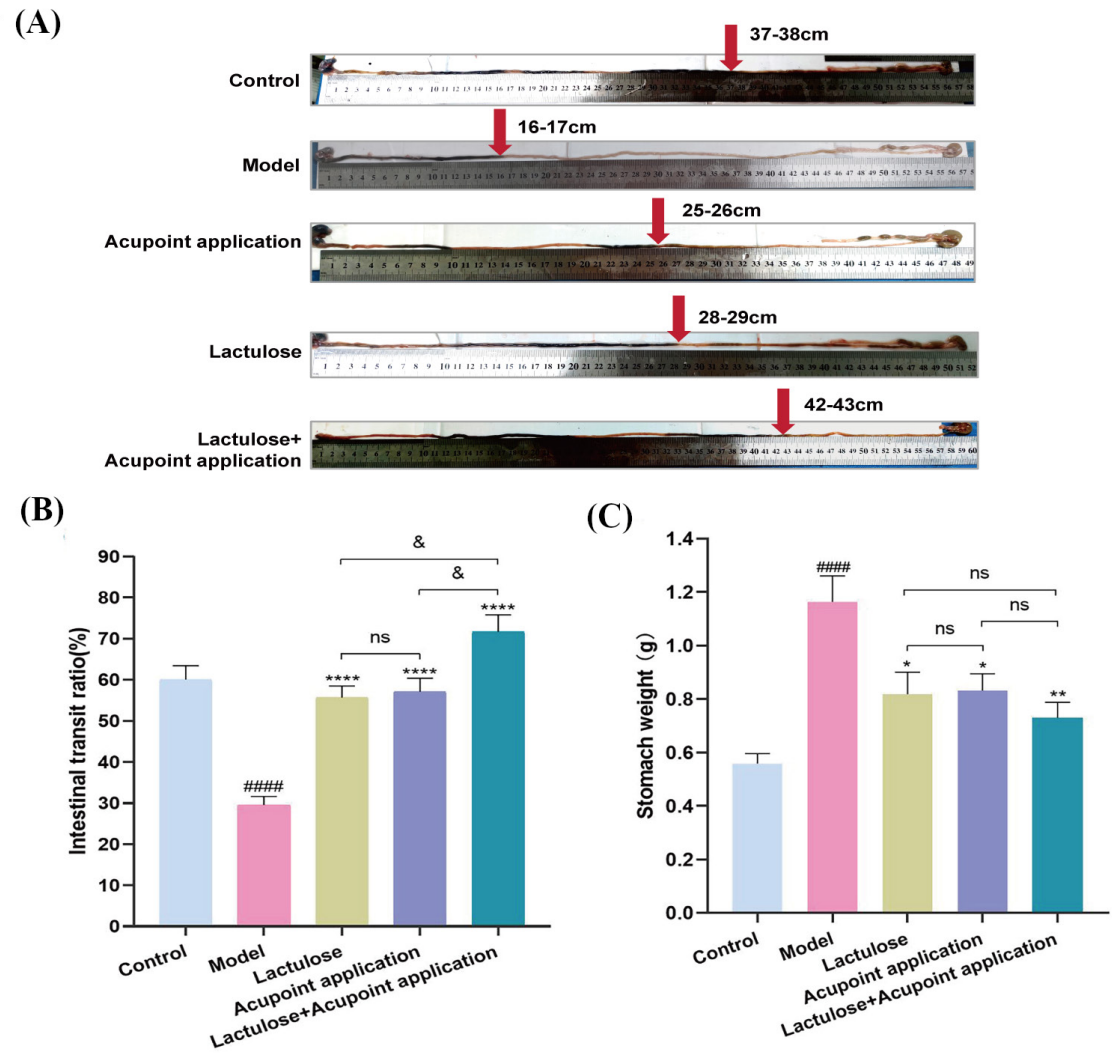

Figure 2. Effects of acupoint application on small intestinal propulsion in constipated mice. (**A**) The propulsive distance of ink in semi-solid nutrients in the intestinal tract. (**B**) Intestinal transit rates. (**C**) Gastric residue. Data were expressed as the mean ± SEM ($n = 8$). $^{\#\#\#\#}p < 0.0001$ *vs.* the control group; $^{*}p < 0.05$, $^{**}p < 0.01$ and $^{****}p < 0.0001$ *vs.* the model group; $^{\&}p < 0.05$ *vs.* the lactulose + acupoint application group; ns, not significant.

group, and lactulose + acupoint application group improved, and especially the thickness of the muscle and mucous membrane, which was significantly thicker than that in the model group (Figures 3B-3C). Moreover, there were no significant differences in the thickness of colonic muscle and mucosa between the acupoint application group and the lactulose group ($p > 0.05$). In addition, the thickness of the colon muscle and mucosa in the lactulose + acupoint application group increased compared to that in the lactulose group or the acupoint application group, but the difference was not statistically significant ($p > 0.05$). These results indicated that the acupoint application of Tongfu Powder significantly reversed the damage caused by loperamide to colon histology. In addition, the effect of acupoint application on colon recovery is not inferior to lactulose, and the combined therapy showed superior efficacy.

3.3. Effects of acupoint application on the SCF/c-kit signaling pathway

Interstitial cells of Cajal (ICC) can regulate gastrointestinal motility, and SCF/c-kit is essential for the growth and development of ICC (*25*). Therefore, SCF/c-kit signaling pathway-related proteins were examined to evaluate the mechanism of acupoint application of Tongfu Powder in treating constipation. RT-qPCR indicated that acupoint application increased the levels of c-kit and SCF genes in constipated mice (Figure 3D). WB also indicated that, the levels of c-kit and SCF increased in the three treatment groups compared to the model group (Figures 3E-3F). The most significant increase was noted in the lactulose + acupoint application group ($p < 0.0001$ and $p < 0.001$). In addition, there were no significant differences in expression of the c-kit and SCF gene and protein in the acupoint application group compared to those in the lactulose group ($p > 0.05$). These results indicated that acupoint application with Tongfu Powder regulated intestinal motility by modulating the expression of c-kit and SCF in colonic tissue. Moreover, the effect of acupoint application on

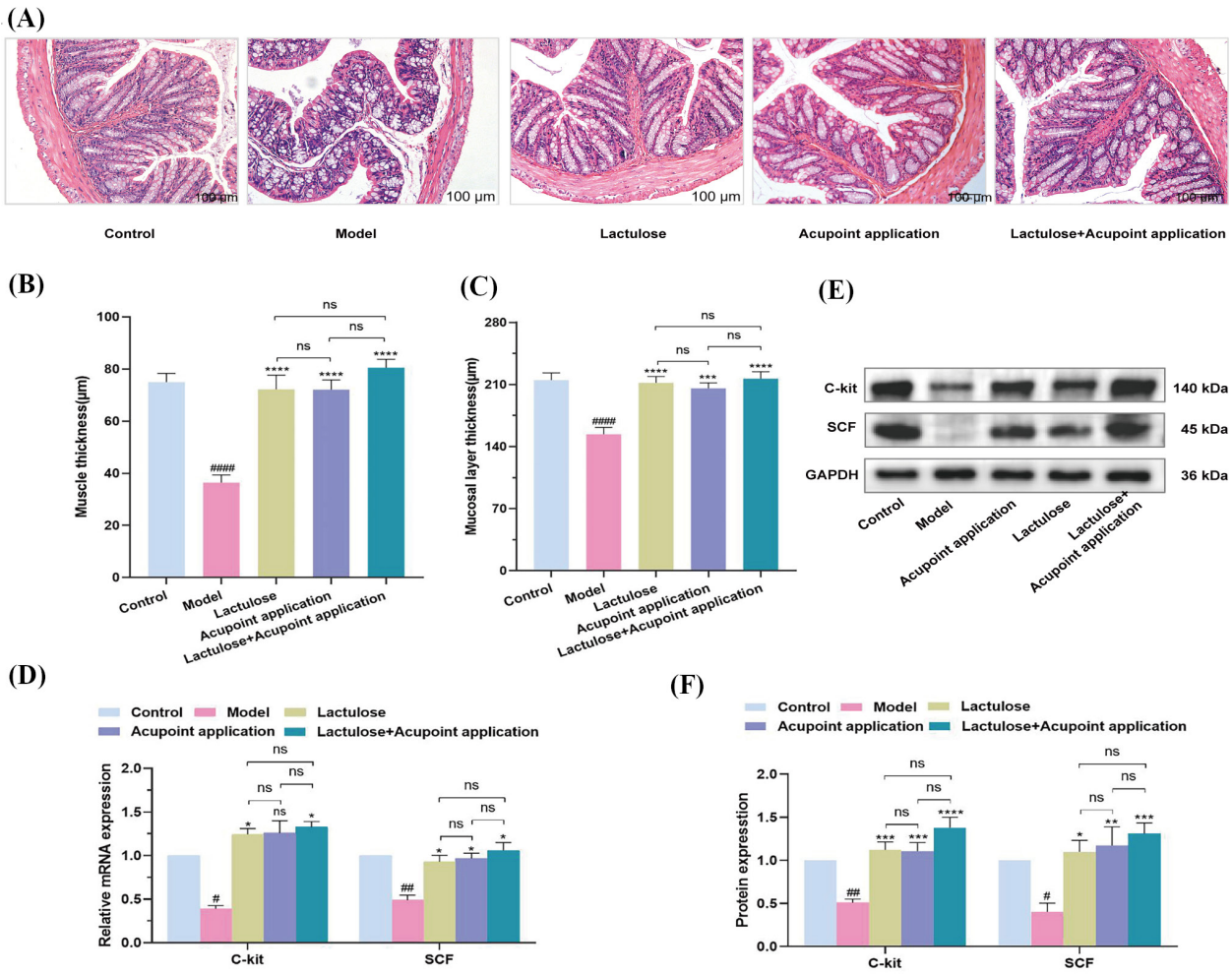

Figure 3. Effects of acupoint application on intestinal tissue damage and the SCF/c-kit signaling pathway in constipated mice. (**A**) Pathological features of colon tissue detected with H&E staining (100× magnification). (**B**) Thickness of colon muscle. (**C**) Thickness of colonic mucosa. (**D**) Expression of c-kit and SCF mRNA detected with RT-qPCR. The levels of C-Kit and SCF protein according to WB (**E**) and quantified using Image J 1.54d (**F**). Data were expressed as the mean ± SEM ($n = 8$). $^{\#}p < 0.05$, $^{\#\#}P < 0.01$, $^{\#\#\#}p < 0.001$, and $^{\#\#\#\#}p < 0.0001$ *vs.* the control group; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ and $^{****}p < 0.0001$ *vs.* the model group; ns, not significant.

colon recovery was not inferior to that of lactulose, and the combination of lactulose and acupoint application appeared to offer additional advantages.

3.4. Effects of acupoint application on the 5-hydroxytryptamine (5-HT) signaling pathway

To further study the mechanism of acupoint application of Tongfu Powder on intestinal peristalsis, 5-HT signaling was analyzed using ELISA and the expression of 5-HT-related proteins was analyzed using WB. As shown in Figure 4, acupoint application of Tongfu Powder significantly increased the serum levels of 5-HT, up-regulated the expression of TPH1 and HTR4, and decreased the expression of SERT compared to levels in the model group. In addition, there were no significant differences in the expression of 5-HT and 5-HT-related proteins between the acupoint application group and the lactulose group ($p > 0.05$). Expression of 5-HT and 5-HT-related proteins tended to improve in the lactulose + acupoint application group compared to the lactulose group or acupoint application group, but the difference was not statistically significant ($p > 0.05$). These results indicate that acupoint application of Tongfu Powder regulated intestinal motility *via* modulation of the 5-HT signaling pathway. Furthermore, the regulatory effect of acupoint application on colonic function was not inferior to that of lactulose, and the combined treatment was superior to either monotherapy alone.

3.5. Effects of acupoint application on the function of the colonic mucus barrier

The colonic mucus layer is reported to play a crucial role in alleviating constipation (*26*). Therefore, the content of the acidic mucus layer in the colon was detected using AB staining (Figure 5A). The quantified image of the blue stained area in Figure 5B indicates that, in comparison to the control group, the blue stained area of the colonic mucus diminished significantly in the model group. However, the blue stained area increased significantly in the lactulose group, acupoint application group, and lactulose + acupoint application group, with the most pronounced increase in the lactulose + acupoint application group. In addition, the level of mucin2 (MUC2), which is involved in the production of colonic mucus, was also examined. Findings revealed that the immunohistochemical detection and image quantification of MUC2 exhibited a similar trend as AB staining (Figures 5C-5D), which further validated the results of

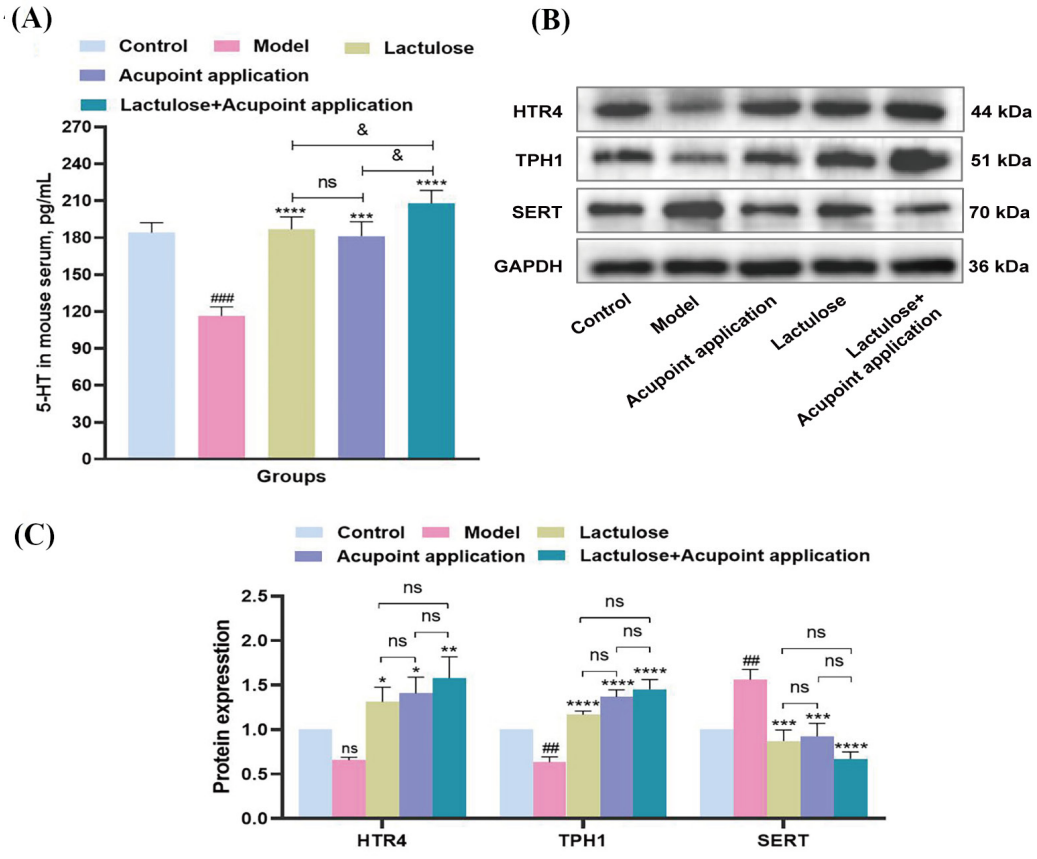

Figure 4. Effects of acupoint application on the 5-hydroxytryptamine (5-HT) signaling pathway. (**A**) The levels of 5-HT in serum measured with ELISA. The levels of HTR4, TPH1, and SERT protein according to WB (**B**) and quantified using Image J 1.48 (**C**). Data were expressed as the mean ± SEM ($n = 6$). $^{\#\#}P < 0.01$, and $^{\#\#\#}p < 0.001$ *vs.* the control group; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ and $^{****}p < 0.0001$ *vs.* the model group; $^{\&}p < 0.05$ *vs.* the lactulose + acupoint application group; ns, not significant.

AB staining.

In addition, there were no significant differences in the blue stained area and the proportion of MUC2 expression in the acupoint application group compared to the lactulose group ($p > 0.05$). Compared to the acupoint application group and the lactulose group, the proportion of the blue stained area in the lactulose + acupoint application group increased significantly ($p < 0.05$). These results indicate that acupoint application of Tongfu Powder improved the dry and hard feces of constipated mice by increasing the production of colonic mucus, that the increase in colonic mucus as a result of acupoint application was not inferior to lactulose, and that the combined treatment has advantages.

3.6. Effects of acupoint application on the function of the colonic epithelial barrier

To investigate the impact of acupoint application of Tongfu Powder on the colonic epithelial barrier in constipated mice, the expression of mRNA of the tight junction proteins Occludin and Claudin-1 was analyzed using RT-qPCR and WB. After acupoint application of Tongfu Powder, the expression of Claudin-1 mRNA and protein increased significantly, and the increase in Claudin-1 was more pronounced in the lactulose + acupoint application group ($p < 0.05$, Figures 5E-5G). Moreover, the expression of Occludin mRNA and protein in the treatment groups was up-regulated compared to

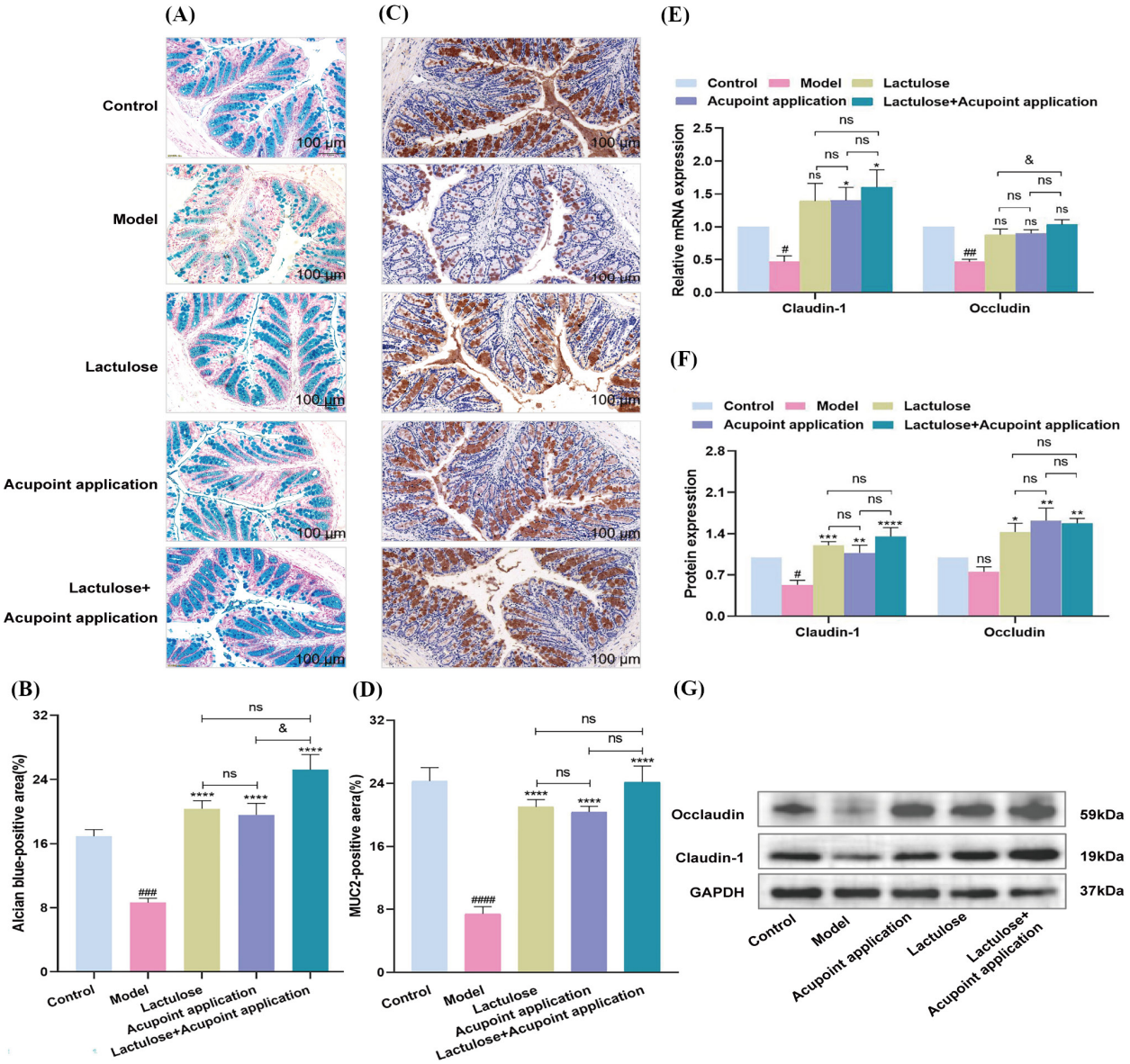

Figure 5. Effects of acupoint application on the function of the colonic mucus and epithelial barriers. The amount of colonic mucus evaluated with AB staining (100× magnification) (**A**) and quantified using Image J 1.48 (**B**). The levels of MUC2 protein detected with IHC staining (100× magnification) (**C**) and quantified using Image J 1.48 (**D**). (**E**) Expression of Occludin and Claudin-1 mRNA detected with RT-qPCR. The levels of Occludin and Claudin-1 protein according to WB (**G**) and quantified using Image J 1.48 (**F**). Data were expressed as the mean ± SEM ($n = 6$). $^{\#}P < 0.05$, $^{\#\#}p < 0.01$, $^{\#\#\#}p < 0.001$, and $^{\#\#\#\#}p < 0.0001$ *vs.* the control group; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ and $^{****}p < 0.0001$ *vs.* the model group; $^{\&}p < 0.05$ *vs.* the lactulose + acupoint application group; ns, not significant.

that in the model group ($p < 0.05$). This suggested that acupoint application of Tongfu Powder restored the barrier function of colon tissue by up-regulating the expression of the connective proteins Claudin-1 and Occludin and that the combined effect is superior to either treatment alone.

3.7. Effects of acupoint application of Tongfu Powder on the abundance and diversity of the intestinal microbiota in constipated mice

To explore the potential role of the intestinal microbiota in the treatment of constipation by acupoint application of Tongfu Powder, the V3-V4 region of the 16S rDNA gene was sequenced. Results yielded 2,080,127 pieces of original data, and the effective data after quality control accounted for 100% of the original data, indicating that sequencing was highly effective. After sequencing, the final ASV feature table and feature sequence were analyzed and annotated. By comparing the bacterial sequences (Figure 6A), one can intuitively see the number of common and unique ASVs in each group. A point worth noting is that the number of ASVs in the model group was significantly smaller than that in other groups. The α diversity reflects the intra-group diversity of microbial samples, and the Goods coverage dilution curve (rarefaction curve) indicates the rate of microbial coverage and changes in species. When the curves of all samples reached the plateau period and the value approached one, this indicated that the sequencing quantity was reasonable and sufficient (Figure 6B). The Chao1 index and observed species index reflected the number of microbial species (Figures 6C-6D). The Shannon index and Simpson index were used to estimate the species diversity in the microbial sample group (Figures 6E-6F). Results indicated that the number and diversity of microbial species in the model group decreased significantly ($p < 0.01$, compared to the control group), while the change in the Alpha diversity index caused by loperamide was effectively reversed in the acupoint application group ($p < 0.01$, compared to the model group).

The β diversity explains the inter-group diversity of microbial samples, which was mainly evaluated by principal component analysis (PCA, Figure 6G), principal coordinates analysis (PCoA, Figure 6H), and clustering analysis (UPGMA, Figure 6I). Results indicated that the similarity within each group and the difference between groups was obvious, and the model group and control group were relatively separate. The three treatment groups were close to the control group. These results indicate that acupoint application of Tongfu Powder has a certain restorative effect on the decrease in intestinal microbiota abundance and diversity in mice with constipation induced by loperamide.

3.8. Effects of acupoint application on reconfiguring the composition of the gut microbiota in constipated mice

The impact of acupoint application on the composition of the intestinal microbiota in mice was investigated at the phylum, family, and genus levels. In the species relative abundance cluster diagram (Figures 7A, 7D, 7I), the distance between the control group and the lactulose group, acupoint application group, and lactulose + acupoint application group was similar, as was the abundance of each species, but they obviously differed from the model group.

The top 30 species were analyzed at the phylum and genus levels. At the phylum level, the main bacteria found in mouse feces were Firmicutes, Bacteroidetes, Proteobacteria, and Verrucomicrobiota. At the phylum level, the proportion of Firmicutes, Proteobacteria, and Bacteroideta decreased in the model group compared to the control group, while the proportion of Verrucomicrobiota increased. These changes in the microbiota were reversed in the treatment groups (Figures 7B-7C). At the family level, loperamide reduced the population size of Lachnospiraceae, Muribaculaceae, Prevotellaceae, Bacteroidaceae, Bifidobacteriaceae, Family_XI, and Veillonellaceae (Figures 7E-7G). The population size of Akkermansiaceae increased (Figure 7H). However, recovery of the abundance of these species was evident to varying degrees in the treatment groups. At the genus level, the model group had a significant increase in the abundance of *Akkermansia* compared to the control group, while the abundance of *Alloprevotella*, *Oscillibacter*, *Helicobacter*, *Anaerotruncus*, *Acetatifactor*, *Clostridium*, *Muribaculum*, and *Bacteroides* decreased. In contrast, recovery of their abundance was evident to different degrees in the treatment groups (Figures 7J-7N). The abundance of *Acetatifactor*, *Clostridium*, *Muribaculum*, and *Bacteroides* increased significantly in the lactulose + acupoint application group ($p < 0.01$). Acupoint application may regulate the intestinal microecology by increasing the abundance of potentially beneficial bacteria and reducing the abundance of potentially harmful bacteria, thus promoting gastrointestinal peristalsis.

3.9. Prediction of the function of the intestinal microbiota in constipated mice

PICRUSt2 is a bioinformatic tool designed to predict the functional potential of gut microbiota based on 16S rRNA gene sequencing data. STAMP analysis and the KEGG database were used to select the top 20 statistically significant functions to display in the form of pictures. As shown in Figure 8A, the histogram on the left represents the differences between groups, with different colors indicating different groups. The higher the column, the higher the abundance ratio, while the dot graph on the right shows the abundance of different

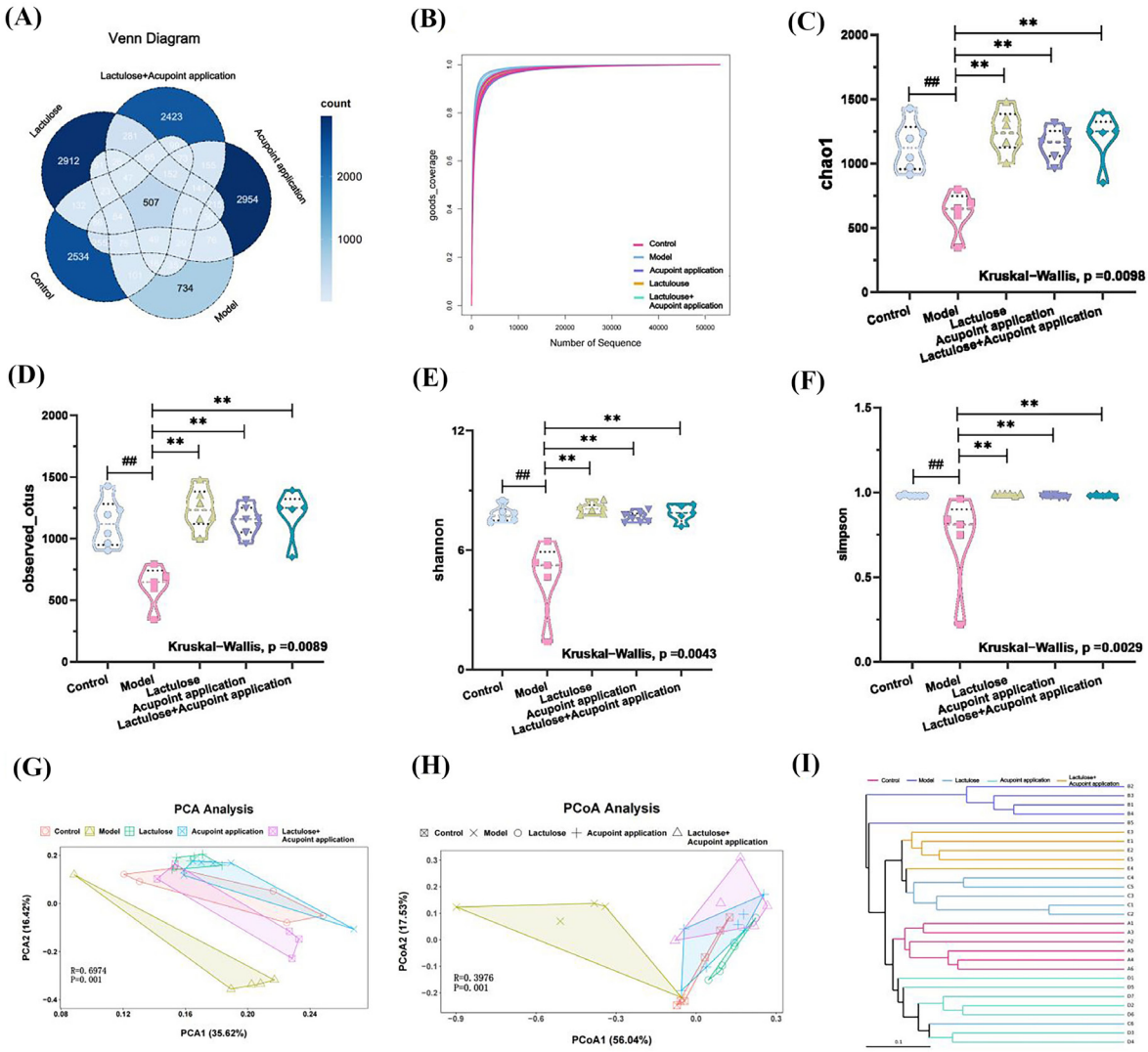

Figure 6. Effects of acupoint application on the abundance and diversity of the gut microbiota in constipated mice. (**A**) Venn diagram of characteristic sequence ASVs. (**B**) The Goods _coverage rarefaction curve. (**C**) The Chao1 index. (D) The observed species index. (**E**) The Shannon index. (**F**) The Simpson index. (**G**) PCA analysis. (**H**) PCoA analysis. (**I**) Clustering analysis. Data were expressed as the mean ± SEM (n = 5-7). $^{\#\#}p < 0.01$ *vs.* the control group; $^{**}p < 0.01$ *vs.* the model group.

items. If the abundance of an item is higher in the control group than in the model group, the dot will be displayed on the right and marked with the color of the control group. Otherwise, it will be marked on the left with the color of the model group. The middle dotted line represents the 95% confidence interval. The farther a point is from this line, the greater the difference. The right axis shows the p value of the difference, with values increasing from bottom to top. The results of microbial function prediction indicated that the constipation model group had decreased levels of restriction enzymes, taurine and hypotaurine metabolism, primary and secondary bile acid biosynthesis, and limonene and pinene degradation compared to the control group. After the application of Tongfu Powder at the acupoint (Figure 8B), however, the metabolic pathways that were predicted to decrease instead increased. In addition, the signaling pathways were analyzed (Figure 8C), which revealed that acupoint application stimulated mannan degradation, heme biosynthesis, CMP− legion salt biosynthesis I, and L-glutamic acid degradation V while inhibiting pyrimidine deoxynucleotide phosphorylation.

3.10. Prediction of the bacterial phenotypes in constipated mice as a result of acupoint application

The tool BugBase was used to predict the bacterial phenotypes of intestinal microorganisms. The levels of aerobic bacteria, mobile elements, biofilm formation, Gram-negative bacteria, and stress tolerance increased in the model group compared to levels in the control group (Figures 9A, 9C, 9D, 9E, and 9H), while the levels of anaerobic bacteria, Gram-positive bacteria, and potentially pathogenic bacteria tended to decrease (Figure 9B, 9F, and 9G). These parameters returned to normal in the lactulose group, acupoint application group, and

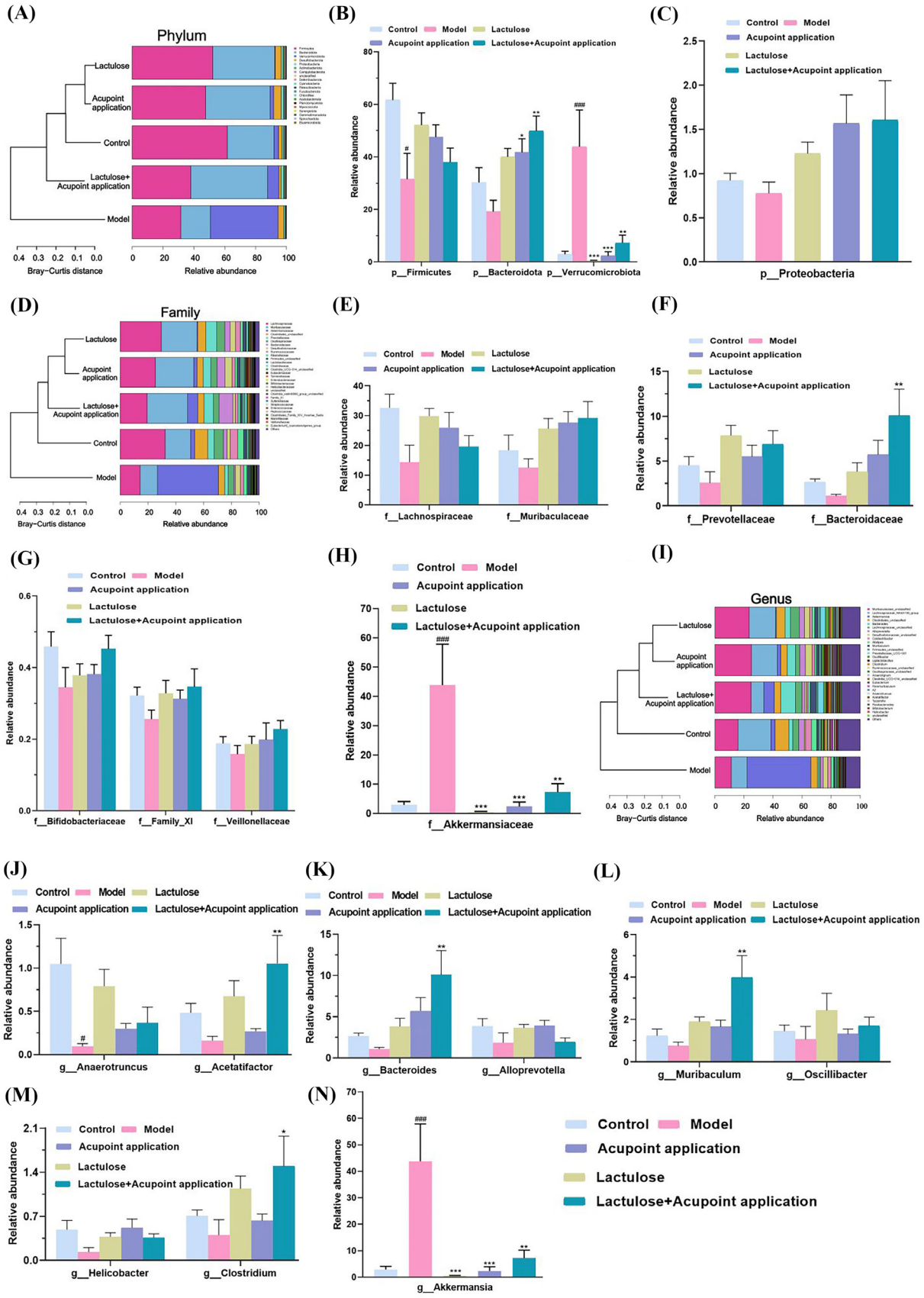

Figure 7. Effects of acupoint application on reconfiguring the composition of the gut microbiota in constipated mice. The differences in bacteria among each group are distinguished by a cluster diagram and histogram at the phylum (**A**), family (**D**), and genus (**I**) levels. (**B-C**) Relative abundance of the gut microbiota at the phylum, level. (**E-H**) Relative abundance of the gut microbiota at the family level. (**J-N**) Relative abundance of the gut microbiota at the genus level. Data were expressed as the mean ± SEM (n = 5-7). $^{\#}p < 0.05$, and $^{\#\#\#}p < 0.001$ *vs.* the control group; $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$ *vs.* the model group.

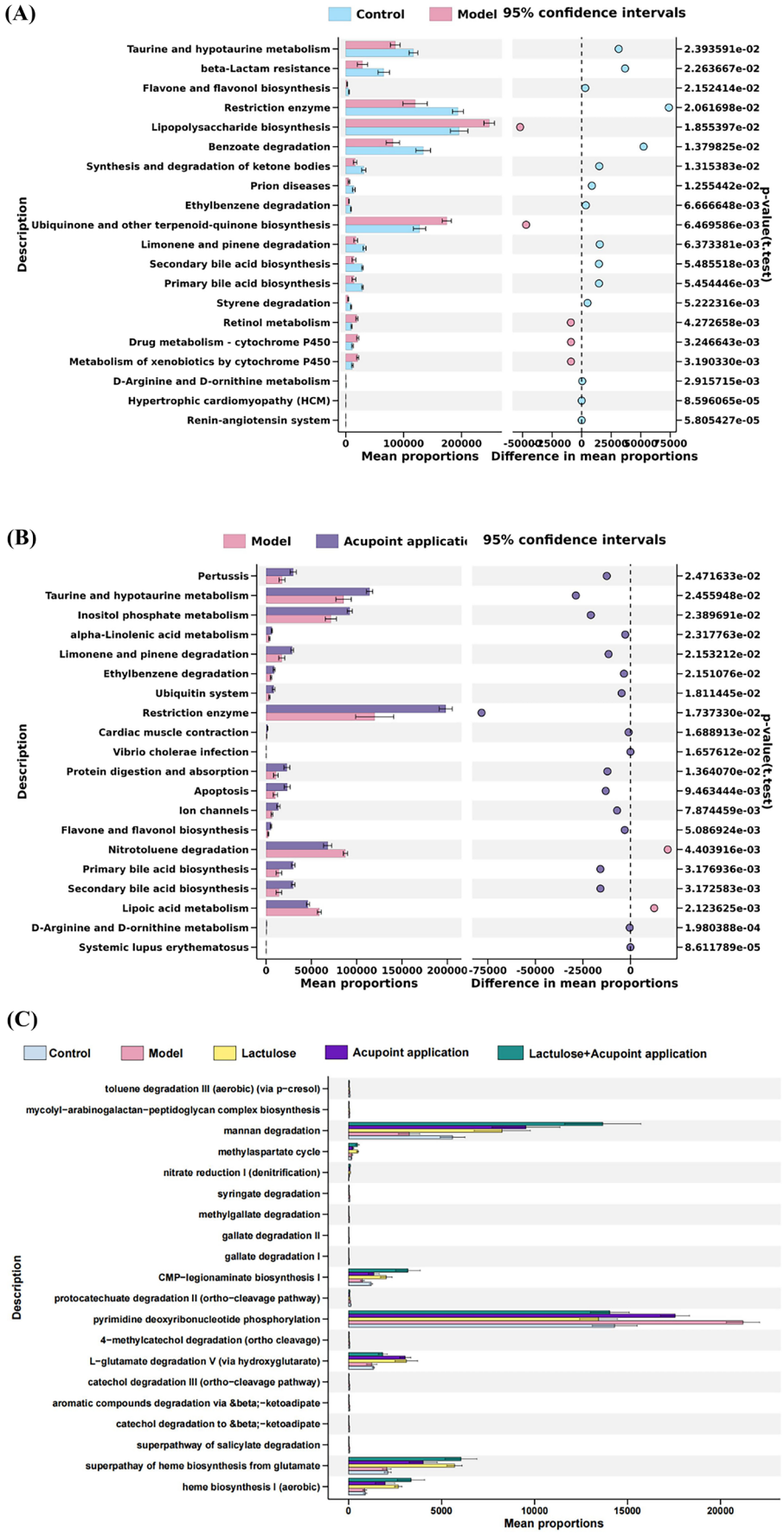

Figure 8. Prediction of the function of acupoint application in intestinal microbiota in constipated mice (*n* = 5-7). (**A**) Comparison between the control group and model group. (**B**) Comparison between the acupoint application group and model group. (**C**) Prediction of signal path function. STAMP analysis and the KEGG database were used to predict PICRUSt function, and the difference was analyzed when $p < 0.05$.

lactulose + acupoint application group. A point worth noting is that the number of beneficial bacteria decreased in the model group, while it increased in the three treatment groups. This phenomenon may be related to the significant increase in Akkermansia spp., which is widely considered a beneficial microbiota, in the model group. These results confirmed that the microbiota plays an important role in the treatment of constipation through acupoint application of Tongfu Powder, and they provide a new direction for future research.

4. Discussion

Acupoint application of Chinese herbal medicines, as one of the characteristic forms of Chinese medicine treatment, has been widely used clinically. Clinical studies demonstrated that applying *Rheum officinale* Baill. to the Shenque acupoint significantly improved the fecal characteristics and increased the frequency of defecation within 24 hours in patients with chronic constipation (*27*). A meta-analysis revealed that acupoint application of Chinese herbal medicines not only relieves constipation symptoms, shortens defecation time, promotes spontaneous complete defecation, and reduces rates of recurrence but also enhances the quality of life for patients (*28*). The current study explored the potential mechanisms of acupoint application of Tongfu Powder in treating chronic constipation in the context of the intestinal barrier and microbiota.

In this study, loperamide-induced constipated mice were used to investigate the alleviation of constipation by acupoint application. After 21 days of loperamide, mice exhibited a decrease in body weight, food intake, fecal water content, and a longer colonic transit time. However, treatment with acupoint application or lactulose significantly improved the overall defecation of constipated mice, including an increase in the number of fecal pellets, fecal weight, and water content, a decrease in gastric residue, and a significant increase in the intestinal propulsion rate (Figures 1-2). Acupoint application or lactulose also alleviated loperamide-induced colonic inflammation and histopathological deterioration such as cellular infiltration and thinning of the muscular and mucosal layers in constipated mice (Figure 3). Notably, the group treated with acupoint application and lactulose displayed the most obvious effect. These results suggest that acupoint application significantly restored gastrointestinal motility by increasing fecal moisture, promoting colonic transit, and alleviating colonic inflammation. Notably, the combined therapy of acupoint application and lactulose conferred additional advantages.

As pacemaker cells of the gastrointestinal tract, ICC can regulate gastrointestinal motility (*29*). A decrease in or damage to ICC in the gastrointestinal tract promotes the development and progression of constipation (*30*). c-kit is a specific marker of ICC. The combination of c-kit and its ligand SCF play an important role in maintaining

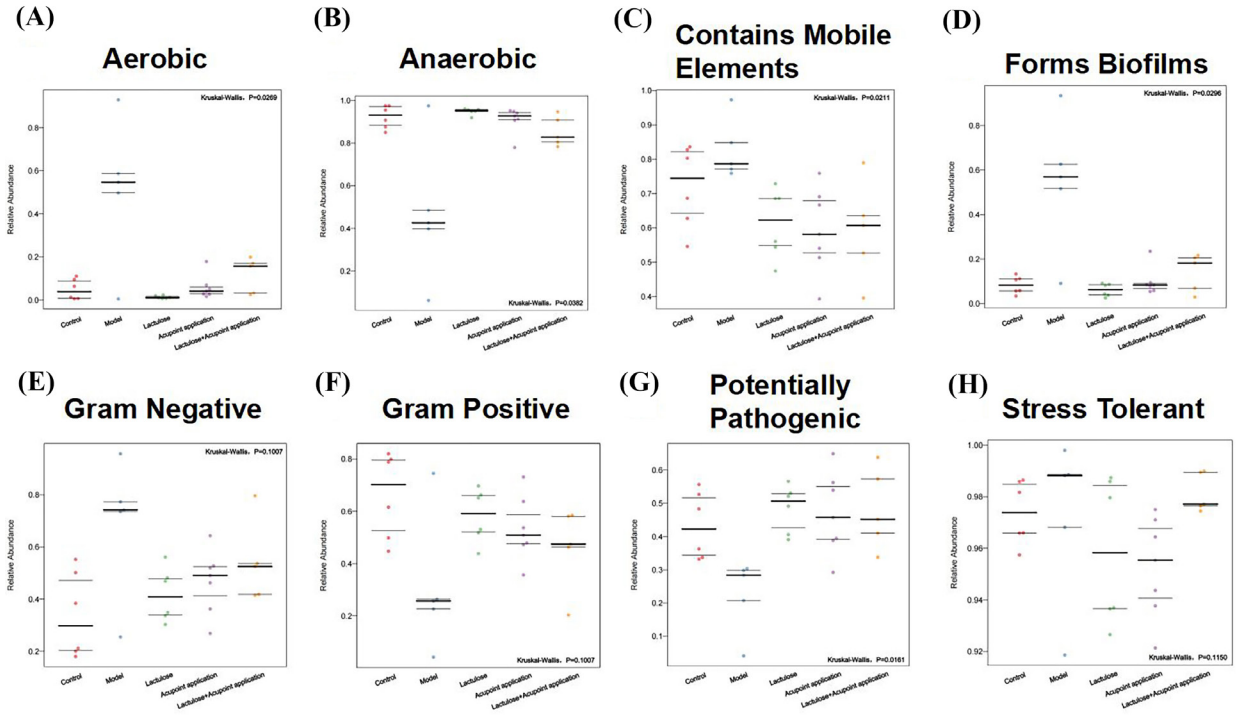

Figure 9. Prediction of the bacterial phenotypes of the intestinal microbiota in constipated mice as a result of acupoint application (*n* = 5-7). (**A**) Aerobic. (**B**) Anaerobic. (**C**) Contains Mobile Elements. (**D**) Forms Biofilms. (**E**) Gram-negative. (**F**) Gram-positive. (**G**) Potentially Pathogenic. (**H**) Stress Tolerant.

the differentiation, development, and maintenance of ICC (*25*). Damage to ICC can be reduced by up-regulating the levels of SCF and c-kit in colon tissue, thereby improving the contractility of colon tissue (*31*). The current results indicated that the levels of c-kit and SCF protein expression were up-regulated in the intestinal tissues of the acupoint application and lactulose treatment groups (Figure 3). Notably, the group treated with acupoint application and lactulose displayed the most obvious effect. Therefore, acupoint application might alleviate constipation by up-regulating the c-kit/SCF signaling pathway to increase the number of ICC, thus enhancing gastrointestinal motility.

5-Hydroxytryptamine (5-HT) is a brain neurotransmitter mostly secreted by enterochromaffin (EC) cells (*32*). It is an important sensor in gastrointestinal motility. The synthesis and release of 5-HT can enhance gastrointestinal peristalsis and alleviate constipation. 5-HTR4 as a receptor of 5-HT, is expressed in the colonic mucosa, and when activated, promotes propulsive motility and attenuates visceral hypersensitivity (*33*). The current results indicated that 5-HT and 5-HTR4 were significantly stimulated especially by the combination of acupoint application and lactulose, which played an important role in the improvement of intestinal motility (Figure 4). Tryptophan hydroxylase 1 (TPH-1) is a member of the amino acid hydroxylase family and a rate-limiting enzyme for the synthesis of 5-HT. TPH-1 is specific and is present in low levels in tissues, so it is usually used as a specific marker of 5-HT neurons and a differentiating feature of 5-HT neurons (*34*). As illustrated in Figure 4, combined therapy with acupoint application and lactulose exhibited superior advantages, which can effectively promote the expression of TPH-1 and play an important role in the synthesis of 5-HT. Serotonin transporter (SERT) is a trans-membrane transporter with a high affinity for 5-HT, which re-uptakes excessive 5-HT to terminate its physiological effects and is involved in regulating gastrointestinal motility (*35*). The current results indicated that the abnormal increase in SERT expression in the model group may have caused a large amount of 5-HT transfer, resulting in poor intestinal motility. However, SERT expression may be inhibited after acupoint application or/and lactulose, and the accumulation of 5-HT near the intestine may have increased, thus alleviated the poor intestinal motility of constipated mice. On the whole, acupoint application restores the 5-HT signaling pathway by regulating the process of 5-HT synthesis, binding to receptors, and inactivation, thereby alleviating constipation symptoms. Therefore, the 5-HT signaling pathway may be an important target for acupoint application in the treatment of constipation.

The normal operation of the intestine relies not only on its peristalsis ability but also on the complete intestinal mucosal barriers. The barriers consist of the mucus barrier, epithelial barrier, and biological barrier (*i.e.*, gut microbiota). Results suggested that the colonic mucosal barrier is closely related to intestinal microbiota. Acupoint application can alleviate functional constipation by promoting intestinal peristalsis, increasing mucus secretion, and regulating the abundance and composition of the intestinal microbiota.

The mucus barrier is a layer of mucus that covers the outside of intestinal cells, and its main component is mucin 2 (MUC2). The intestinal mucus layer is synthesized and secreted by goblet cells, which not only maintain intestinal homeostasis but also play a role in lubricating the intestine (*36*). When the intestinal mucus layer decreases, it can lead to a large number of harmful substances and bacteria invading the intestinal crypt and intestinal epithelium, causing various intestinal diseases. It can also make the surface of stool dry and difficult to discharge. In order to verify the effect of acupoint application on the mucus barrier, the level of colonic acidic mucus and mucin MUC2 was measured. As shown in Figure 5, the stained area of the colonic mucus layer and the level of MUC2 protein increased significantly after acupoint application or lactulose treatment compared to that in the model group. Notably, the group treated with acupoint application and lactulose displayed the most obvious effect. These results suggest that acupoint application greatly increased the quantity of acidic mucus on the surface of the colonic mucosa, thus alleviating the symptoms of dry stool in mice with constipation.

The complete and continuous epithelial barrier is also the basic intestinal mucosal barrier that protects the colon. Tight junctions are the core structure of the epithelial barrier, in which Occludin, Claudin, and other proteins together constitute the main body of tight junctions. The integrity of the tight junction structure is conducive to the closure of intercellular space and plays an important role in regulating cell permeability and preventing harmful substances from invading the intestinal cavity (*37*). The current results indicated that the tight junction barrier of the intestinal mucosa was significantly damaged in the model group with the decreased expression of Occludin and Claudin-1 (Figure 5). However, combined treatment with acupoint application and lactulose increased the expression of Occludin and Claudin-1 mRNA and protein in the colon tissue of constipated mice, repaired the tight junction, and restored the intestinal mucosal barrier. Acupoint application combined with lactulose is superior to monotherapy in promoting defecation function and reducing colonic tissue injury in constipated mice.

The intestinal microbiota is another important component of the intestinal mucosal barrier. Mounting evidence indicates that constipation is closely related to a disorder of the intestinal microbiota (*38*). As shown in Figure 6, Chao1, observed species, and the Shannon and Simpson indexes showed that acupoint application reversed the decrease in the number and abundance of

intestinal microbial communities in constipated mice. Therefore, the mechanism of acupoint application in the treatment of constipation is also inseparable from the regulation of the intestinal microbiota.

To further study the effect of acupoint application on the bacterial community at the phylum level, the differences in the composition of the bacterial community were analyzed. Short-chain fatty acids (SCFAs) are intestinal microbial metabolites that promote intestinal motility and relieve constipation (*39*). In the current study, Lachnospiraceae, Bifidobacteriaceae, and Prevotellaceae decreased significantly in the model group compared to the control group, and these are common bacteria that produce SCFAs. The difficulty in defecation in the model group may be related to the reduction in SCFA-producing bacteria. Fortunately, the abundance of these bacteria increased to varying degrees after acupoint application or lactulose (Figure 7). Moreover, the effect of acupoint application combined with lactulose is superior to either treatment alone.

Akkermansia muciniphila (Akk) is a mucin-degrading bacterium belonging to the phylum Verrucomicrobiota. It colonizes the mucous layer of the gastrointestinal tract and has the ability to degrade mucin. When Akk grows excessively, it may cause excessive degradation of the mucin layer in the intestine, leading to an increase in intestinal permeability and leakage. This can increase the likelihood of toxins, bacterial lipopolysaccharides, and even pathogenic bacteria entering the bloodstream and increasing the risk of intestinal diseases (*40*). In an experiment involving 16S sequencing of fecal samples from constipated mice, researchers found a correlation between proinflammatory metabolites and Akk (*41*). This may be related to the destruction of the mucus barrier due to excessive Akk degradation of mucus protein in the feces of constipated mice. In the current study, there was a significant increase in Akk in the intestinal microbiota of constipated mice but a significant decrease after acupoint application of Tongfu Powder and/or lactulose treatment ($p < 0.01$). This finding is consistent with the results of previous research (*42,43*), suggesting that acupoint application may alter the composition and metabolism of the intestinal microbiota by regulating the number of Akk, thus alleviating constipation symptoms. However, some studies have shown that Akk is a beneficial bacterium in the intestine that helps maintain intestinal homeostasis (*44,45*). This seems to contradict the current results. In order to solve this mystery, further research is needed to reveal the complex role of Akk in the intestine and its relationship to constipation and other diseases. In other words, Akk, as a unique mucus-degrading bacterium, plays an important role in intestinal health. It participates in maintaining intestinal homeostasis by degrading mucin, but it may also contribute to intestinal diseases under certain conditions. Therefore, the biological characteristics of Akk and its

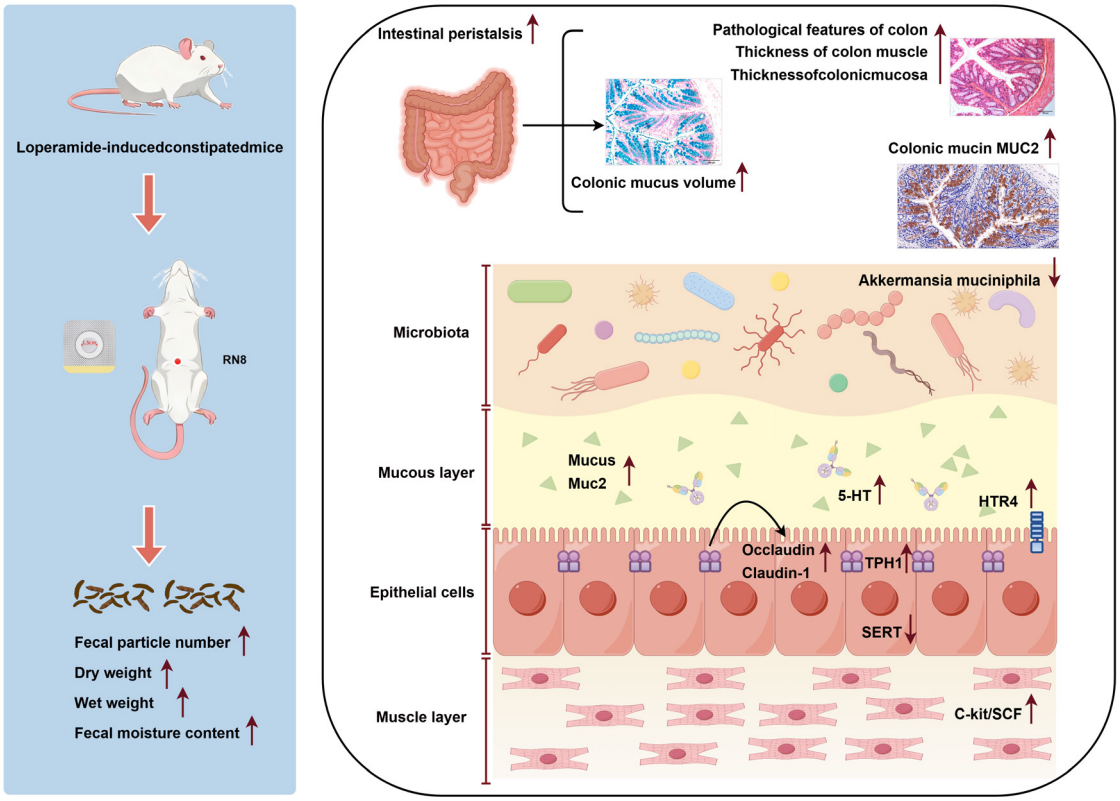

Figure 10. Acupoint application of Tongfu Powder might alleviate loperamide-induced constipation by regulating the intestinal barrier and gut microbiota.

mechanism of action in the intestine need to be studied further in order to provide new ideas and methods for preventing and treating intestinal diseases.

Apart from analyzing the composition of species in the intestinal microbiota, this study also predicted the function and bacterial phenotypes of the intestinal microbiota. As shown in Figure 8 and Figure 9, restriction enzymes, taurine and hypotaurine metabolism, primary and secondary bile acid biosynthesis, and limonene and pinene degradation are the main factors affecting the ecological imbalance of the intestinal microbiota. At the same time, acupoint application may play a role by stimulating mannan degradation, heme biosynthesis, CMP-legionaminate acid biosynthesis I, and L-glutamic acid degradation V signal pathways and inhibiting pyrimidine deoxynucleotide phosphorylation. In addition, the prediction of bacterial phenotypes revealed that the increase in aerobic bacteria, mobile genetic elements, biofilm-forming bacteria, Gram-negative bacteria, and stress-tolerant bacteria and the decrease in anaerobic bacteria and Gram-positive bacteria are related to the development of constipation. A point worth noting is that mobile genetic elements and biofilm-forming bacteria have been widely studied in many intestinal diseases due to their potential pathogenicity and their impact on the human immune system. The presence of these characteristics not only enhances the adaptability of bacteria to the external environment but also mediates their toxicity to the host, thus exacerbating constipation symptoms (*46*). However, the finding that acupoint application can reverse these abnormal bacterial phenotypes and restore them to normal levels is encouraging.

5. Conclusions

In summary, the current study demonstrated that acupoint application of Tongfu Powder acted to counter constipation as it effectively restored gastrointestinal motility by increasing fecal moisture, promoting colonic transit, and relieving colonic inflammation. The potential mechanisms of acupoint application might be related to the regulation of the intestinal barrier and gut microbiota. It significantly increased the number of ICC and enhanced intestinal 5-HT synthesis. Acupoint application of Tongfu Powder also promoted MUC2 secretion, increased the expression of tight junction proteins (Claudin-1 and Occludin), and improved the composition of the gut microbiota (Figure 10). In addition, the laxative effect of Tongfu Powder acupoint application was not inferior to that of lactulose, and the combination therapy produced a superior effect. Therefore, this study provides a reliable theoretical foundation for the clinical application of acupoint application, which may serve as a promising alternative therapy for constipation.

Funding: This study was funded by the Shandong Province Plan for Scientific and Technological Development of Traditional Chinese Medicine (Grant No. M20243602), The Sixth Batch of National Outstanding Clinical Talents in Traditional Chinese Medicine Training Program by the State Administration of Traditional Chinese Medicine (SATCM Document on Personnel and Education No. 256 [2025]), and 2024 Qilu Bian-cang Traditional Chinese Medicine Talent Cultivation Project.

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Aziz I, Whitehead WE, Palsson OS, Törnblom H, Simrén M. An approach to the diagnosis and management of Rome IV functional disorders of chronic constipation. Expert Rev Gastroenterol Hepatol. 2020; 14:39-46.
2. Ihara E, Manabe N, Ohkubo H, *et al*. Evidence-Based Clinical Guidelines for Chronic Constipation 2023. Digestion. 2025; 106:62-89.
3. Barbara G, Barbaro MR, Marasco G, Cremon C. Chronic constipation: from pathophysiology to management. Minerva Gastroenterol (Torino). 2023; 69:277-290.
4. Harris LA, Chang CH. Burden of Constipation: Looking Beyond Bowel Movements. Am J Gastroenterol. 2022; 117:S2-S5.
5. Chang L, Chey WD, Imdad A, Almario CV, Bharucha AE, Diem S, Greer KB, Hanson B, Harris LA, Ko C, Murad MH, Patel A, Shah ED, Lembo AJ, Sultan S. American Gastroenterological Association-American College of Gastroenterology Clinical Practice Guideline: Pharmacological Management of Chronic Idiopathic Constipation. Gastroenterology. 2023; 164:1086-1106.
6. Shah ED, Staller K, Nee J, Ahuja NK, Chan WW, Lembo A, Brenner DM, Siegel CA, Chey WD. Evaluating the Impact of Cost on the Treatment Algorithm for Chronic Idiopathic Constipation: Cost-Effectiveness Analysis. Am J Gastroenterol. 2021; 116:2118-2127.
7. Tan S, Peng C, Lin X, Peng C, Yang Y, Liu S, Huang L, Bian Y, Li Y, Xu C. Clinical efficacy of non-pharmacological treatment of functional constipation: a systematic review and network meta-analysis. Front Cell Infect Microbiol. 2025; 15:1565801.
8. Cui J, Wang J, Wang Y, Zhang C, Hu G, Wang Z. External treatment of traditional Chinese medicine for functional dyspepsia in children: Protocol for a systematic review and network meta-analysis. Medicine (Baltimore). 2022; 101:e31597.
9. Wang K, Qiu H, Chen F, Cai P, Qi F. Considering traditional Chinese medicine as adjunct therapy in the management of chronic constipation by regulating intestinal flora. Biosci Trends. 2024; 18:127-140.
10. Wang Q, Zhao L, Liu J, Chen L, Zhang B, Zhang Q, Lu Y, Gao Y, Zheng X, He Z, Jing S. Meta analysis of clinical efficacy of acupoint application in the treatment of irritable bowel syndrome. Afr Health Sci. 2024; 24:351-361.
11. Gao H, He C, Xin S, Hua R, Du Y, Wang B, Gong F, Yu X, Pan L, Liang C, Gao L, Shang H, Xu JD. Rhubarb extract rebuilding the mucus homeostasis and regulating mucin-associated flora to relieve constipation. Exp Biol Med

(Maywood). 2023; 248:2449-2463.
12. Wang L, Wu F, Hong Y, Shen L, Zhao L, Lin X. Research progress in the treatment of slow transit constipation by traditional Chinese medicine. J Ethnopharmacol. 2022; 290:115075.
13. Tian H, Huang D, Li T, Huang L, Zheng X, Tang D, Zhang L, Wang J. The protective effects of total phenols in magnolia officinalix rehd. et wils on gastrointestinal tract dysmotility is mainly based on its influence on interstitial cells of cajal. Int J Clin Exp Med. 2015; 8:20279-20286.
14. Li CB, Yang X, Tang WB, Liu CY, Xie DP. Arecoline excites the contraction of distal colonic smooth muscle strips in rats *via* the M3 receptor-extracellular Ca^{2+} influx-Ca^{2+} store release pathway. Can J Physiol Pharmacol. 2010; 88:439-447.
15. Xu M, Wang W, Su S, Li W, Hu X, Zhang J. Arecoline alleviated loperamide induced constipation by regulating gut microbes and the expression of colonic genome. Ecotoxicol Environ Saf. 2023; 264:115423.
16. Kulkarni M, Sawant N, Kolapkar A, Huprikar A, Desai N. Borneol: a Promising Monoterpenoid in Enhancing Drug Delivery Across Various Physiological Barriers. AAPS PharmSciTech. 2021; 22:145.
17. Zheng H, Chen Y, Lu S, Liu Z, Ma Y, Zhang C, Zhang Y, Zhang J, Liu C, Chu M, Pei F, Liu S, Duan L. Mechanosensory Piezo2 regulated by gut microbiota participates in the development of visceral hypersensitivity and intestinal dysmotility. Gut Microbes. 2025; 17:2497399.
18. Xu X, Wang Y, Long Y, Cheng Y. Chronic constipation and gut microbiota: current research insights and therapeutic implications. Postgrad Med J. 2024; 100:890-897.
19. Yuan F, Lu F, Guo Y, Zhang C. Acupoint Application Combined with Acupoint Massage for Treating Constipation in a Patient with Chronic Obstructive Pulmonary Disease. J Vis Exp. 2023; 198.
20. Kojima R, Doihara H, Nozawa K, Kawabata-Shoda E, Yokoyama T, Ito H. Characterization of two models of drug-induced constipation in mice and evaluation of mustard oil in these models. Pharmacology. 2009; 84:227-233.
21. Qi FH, Cai PP, Liu X, Si GM. Adenovirus-mediated P311 ameliorates renal fibrosis through inhibition of epithelial-mesenchymal transition *via* TGF-β1-Smad-ILK pathway in unilateral ureteral obstruction rats. Int J Mol Med. 2018; 41:3015-3023.
22. Zhao L, Cheng N, Sun B, Wang S, Li A, Wang Z, Wang Y, Qi F. Regulatory effects of Ningdong granule on microglia-mediated neuroinflammation in a rat model of Tourette's syndrome. Biosci Trends. 2020; 14:271-278.
23. Qi F, Wang J, Zhao L, Cai P, Tang W, Wang Z. Cinobufacini inhibits epithelial-mesenchymal transition of human hepatocellular carcinoma cells through c-Met/ERK signaling pathway. Biosci Trends. 2018; 12:291-297.
24. Ma YN, Jiang XM, Hu XQ, Wang L, Gao JJ, Liu H, Qi FH, Song PP, Tang W. Cinobufacini Inhibits Survival and Metastasis of Hepatocellular Carcinoma *via* c-Met Signaling Pathway. Chin J Integr Med. 2025; 31:311-325.
25. Chai Y, Huang Y, Tang H, Tu X, He J, Wang T, Zhang Q, Xiong F, Li D, Qiu Z. Role of stem cell growth factor/c-Kit in the pathogenesis of irritable bowel syndrome. Exp Ther Med. 2017; 13:1187-1193.
26. Chen CM, Wu CC, Huang CL, Chang MY, Cheng SH, Lin CT, Tsai YC. Lactobacillus plantarum PS128 Promotes Intestinal Motility, Mucin Production, and Serotonin Signaling in Mice. Probiotics Antimicrob Proteins. 2022; 14:535-545.
27. Wei L, Luo Y, Zhang X, Liu Y, Gasser M, Tang F, Ouyang WW, Wei H, Lu S, Yang Z, Waaga-Gasser AM, Deng C, Lin M. Topical therapy with rhubarb navel plasters in patients with chronic constipation: Results from a prospective randomized multicenter study. J Ethnopharmacol. 2021; 264:113096.
28. Yan L, Liu H, Yan R, Tan L, Tan J, Lei Y. Effect of traditional Chinese medicine external therapy for functional constipation: a meta-analysis. Am J Transl Res. 2023; 15:13-26.
29. Koh SD, Drumm BT, Lu H, Kim HJ, Ryoo SB, Kim HU, Lee JY, Rhee PL, Wang Q, Gould TW, Heredia D, Perrino BA, Hwang SJ, Ward SM, Sanders KM. Propulsive colonic contractions are mediated by inhibition-driven poststimulus responses that originate in interstitial cells of Cajal. Proc Natl Acad Sci U S A. 2022; 119:e2123020119.
30. Tang X, Huang Y, Jiang T, Wu J, Wang K, Wu W. Pathophysiological mechanisms, diagnostic innovations, and multimodal therapeutic strategies for slow transit constipation. BMC Gastroenterol. 2025; 25:810.
31. Yin J, Liang Y, Wang D, Yan Z, Yin H, Wu D, Su Q. Naringenin induces laxative effects by upregulating the expression levels of c-Kit and SCF, as well as those of aquaporin 3 in mice with loperamide-induced constipation. Int J Mol Med. 2018; 41:649-658.
32. Wei L, Singh R, Ha SE, Martin AM, Jones LA, Jin B, Jorgensen BG, Zogg H, Chervo T, Gottfried-Blackmore A, Nguyen L, Habtezion A, Spencer NJ, Keating DJ, Sanders KM, Ro S. Serotonin Deficiency Is Associated With Delayed Gastric Emptying. Gastroenterology. 2021; 160:2451-2466.e19.
33. Hoffman JM, Tyler K, MacEachern SJ, Balemba OB, Johnson AC, Brooks EM, Zhao H, Swain GM, Moses PL, Galligan JJ, Sharkey KA, Greenwood-Van Meerveld B, Mawe GM. Activation of colonic mucosal 5-HT(4) receptors accelerates propulsive motility and inhibits visceral hypersensitivity. Gastroenterology. 2012; 142:844-854.e4.
34. Yan X, Ma P, Wang W, Zeng W, Li Y, Hou Y, Ye J, Zheng Q, Zhang W, Yao J, Li Y. Piezo knockdown reduces 5-hydroxytryptamine release from enterochromaffin cells and exacerbates intestinal dyskinesia in mice with functional constipation. Int J Mol Med. 2025; 56:178.
35. Cao H, Liu X, An Y, Zhou G, Liu Y, Xu M, Dong W, Wang S, Yan F, Jiang K, Wang B. Dysbiosis contributes to chronic constipation development *via* regulation of serotonin transporter in the intestine. Sci Rep. 2017; 7:10322.
36. Liu Y, Yu Z, Zhu L, Ma S, Luo Y, Liang H, Liu Q, Chen J, Guli S, Chen X. Orchestration of MUC2 - The key regulatory target of gut barrier and homeostasis: A review. Int J Biol Macromol. 2023; 236:123862.
37. Horowitz A, Chanez-Paredes SD, Haest X, Turner JR. Paracellular permeability and tight junction regulation in gut health and disease. Nat Rev Gastroenterol Hepatol. 2023; 20:417-432.
38. Zhang S, Wang R, Li D, Zhao L, Zhu L. Role of gut microbiota in functional constipation. Gastroenterol Rep (Oxf). 2021; 9:392-401.
39. Liu SH, Yang XF, Liang L, Song BB, Song XM, Yang YJ, Alhoot MA. Regulatory mechanisms of the gut microbiota-short chain fatty acids signaling axis in slow transit constipation and progress in multi-target

interventions. Front Microbiol. 2025; 16:1689597.
40. Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA, Wolter M, Pudlo NA, Kitamoto S, Terrapon N, Muller A, Young VB, Henrissat B, Wilmes P, Stappenbeck TS, Núñez G, Martens EC. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. Cell. 2016; 167:1339-1353.e21.
41. Tan R, Dong H, Chen Z, Jin M, Yin J, Li H, Shi D, Shao Y, Wang H, Chen T, Yang D, Li J. Intestinal Microbiota Mediates High-Fructose and High-Fat Diets to Induce Chronic Intestinal Inflammation. Front Cell Infect Microbiol. 2021; 11:654074.
42. Wang L, Chai M, Wang J, Yu Q, Wang G, Zhang H, Zhao J, Chen W. Bifidobacterium longum relieves constipation by regulating the intestinal barrier of mice. Food Funct. 2022; 13:5037-5049.
43. Wei Y, Huang N, Ye X, Liu M, Wei M, Huang Y. The postbiotic of hawthorn-probiotic ameliorating constipation caused by loperamide in elderly mice by regulating intestinal microecology. Front Nutr. 2023; 10:1103463.
44. Grander C, Adolph TE, Wieser V, *et al*. Recovery of ethanol-induced Akkermansia muciniphila depletion ameliorates alcoholic liver disease. Gut. 2018; 67:891-901.
45. Jiang JG, Luo Q, Li SS, Tan TY, Xiong K, Yang T, Xiao TB. Cinnamic acid regulates the intestinal microbiome and short-chain fatty acids to treat slow transit constipation. World J Gastrointest Pharmacol Ther. 2023; 14:4-21.
46. Forster SC, Liu J, Kumar N, Gulliver EL, Gould JA, Escobar-Zepeda A, Mkandawire T, Pike LJ, Shao Y, Stares MD, Browne HP, Neville BA, Lawley TD. Strain-level characterization of broad host range mobile genetic elements transferring antibiotic resistance from the human microbiome. Nat Commun. 2022; 13:1445.

Received February 18, 2026; Revised April 4, 2026; Accepted April 5, 2026.

**Address correspondence to:*
Fanghua Qi, Traditional Chinese Medicine, Shandong Provincial Hospital affiliated to Shandong First Medical University, No. 324 Jingwuweiqi Road, Ji'nan, Shandong, China 250021.
E-mail: qifanghua2006@126.com

Released online in J-STAGE as advance publication April 15, 2026.

Original Article

DOI: 10.5582/ddt.2025.01123

Cytoderm metabolic-labeling SCMLP-TB for pulmonary tuberculosis diagnosis: A preliminary diagnostic accuracy study

Guangyan Liang[1,§], Guiqin Dai[1,§], Xiaorong Hu[1,§], Deliang Liu[1], Zhiqiang Lin[1], Mengru Yang[1,4], Zhuojun He[1], Peifen Chen[1], Yipeng Liu[1], Xinyun Jia[1], Xiafei Dai[1], Pengfei Zhao[1,*], Mingbin Zheng[1,2,*], Yang Zhou[1,3,*], Hongzhou Lu[1,*]

[1] National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital, Southern University of Science and Technology, Shenzhen, China;
[2] The Affiliated Dongguan Songshan Lake Central Hospital, Guangdong Medical University, Dongguan, China;
[3] Molecular Biology Research Center & Center for Medical Genetics, School of Life Sciences, Central South University, Changsha, China;
[4] School of Public Health, Shenzhen University Medical School, Shenzhen University, Shenzhen, Guangdong, China.

SUMMARY: Tuberculosis (TB) remains a significant global health issue. Early diagnosis is crucial, yet current diagnostic technologies are limited by suboptimal sensitivity. Thus, we developed a novel tuberculosis metabolic labeling probe (single cell metabolic labeling probe for tuberculosis, SCMLP-TB) and evaluated its diagnostic performance. In this retrospective study of 70 suspected TB patients, we calculated the sensitivity and specificity of SCMLP-TB and compared it with culture and Xpert MTB/RIF (Xpert) using the final clinical diagnosis as reference standard. Eligible participants were classified as confirmed TB (CT), clinically diagnosed TB (CDxT), or non-TB cases based on the diagnosis criteria for pulmonary tuberculosis (WS 288-2017). Of the 70 participants, 40 (57.0%) were diagnosed with TB, including 30 CT cases and 10 CDxT cases. The overall diagnostic sensitivity and specificity of SCMLP-TB were 97.5% and 96.7%, respectively. Notably, SCMLP-TB identified 10 CDxT cases missed by both culture and Xpert. The overall diagnostic sensitivity of culture and Xpert was 62.5% and 72.5%, respectively, while both showed a specificity of 100.0%, demonstrating that SCMLP-TB was more sensitive than culture and Xpert. Besides, the fluorescence intensity from TB patients was significantly higher than non-TB patients. The fluorescence intensity showed a significant negative correlation with the time to positivity (TTP) of culture, which suggested that SCMLP-TB could also serve as an indicator of bacterial loads in patients' samples. Consequently, SCMLP-TB demonstrated a promising tool for the sensitive and ultra-fast diagnosis of pulmonary TB suspects, particularly for paucibacillary pulmonary TB.

Keywords: Tuberculosis, SCMLP-TB, metabolic-labeling, accuracy, diagnosis

1. Introduction

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (MTB), is the leading global infectious disease in terms of morbidity and mortality (*1,2*). Early and precise diagnosis is crucial for controlling the TB epidemic (*3*). As the most common tool for TB diagnosis, sputum-smear microscopy suffered from considerable training demands, operator-dependent performance, inability to discriminate live from dead bacilli, and relatively low sensitivity (*4,5*). Culture-based assays were the gold standard of TB diagnosis, but hampered by prolonged incubation, stringent lab requirements, limited sensitivity, and contamination risk leading to inconclusive outcomes (*4,6,7*). Xpert MTB/RIF (Xpert), a polymerase chain reaction (PCR)-based test, improved TB detection yet it remained prohibitively expensive and displayed suboptimal sensitivity in paucibacillary samples, particularly in TB/HIV (human immunodeficiency virus) co-infected individuals and extrapulmonary samples (*7-10*). Therefore, developing a rapid, sensitive, specific and accurate method for diagnosing live MTB was essential for effective TB control.

Reportedly, a fluorescein-modified trehalose analog (FITC-Tre) achieved rapid and accurate detection of viable MTB in patient sputum by engaging in glycolipid biosynthesis and inserting into the bacterial cell wall (*11-13*). As the predominant lipids in the mycomembrane of MTB, trehalose monomycolates and trehalose dimycolates were essential for cellular viability (*13-15*). By antigen-85 (Ag85) complex catalysis, these

glycolipids were secured within the cytoderm (*16,17*). However, the detection efficacy of conventional fluorophores like fluorescein isothiocyanate (FITC) was limited by photobleaching and aggregation-induced quenching at high concentrations or in aggregated states (*11,18,19*).

To solve this problem, we designed SCMLP-TB (single cell metabolic labeling probe for tuberculosis), a fluorescein-tagged trehalose analog that exhibited aggregation-induced emission, which served as an exogenous substrate to selectively aggregate at MTB cytoderm and enabled fluorescent labeling and visualization of viable bacteria (*20-22*).

In this study, we retrospectively evaluated the diagnostic accuracy of SCMLP-TB using sputum or bronchoalveolar lavage fluid (BALF) samples for pulmonary TB detection. Additionally, we preliminarily explored the correlation between fluorescence (FL) intensity and bacterial load evidenced by the time to positivity (TTP) of culture. This approach showed great potential for precise MTB diagnosis and real-time bacterial viability assessment.

2. Materials and Methods

2.1. Study design and participants

This was a single-centre, retrospective study conducted at Shenzhen Third People's Hospital, which was designated as the hospital responsible for the treatment of tuberculosis (TB) patients in Shenzhen and its surrounding areas. From May 2023 to October 2024, we collected data on suspected pulmonary TB patients who were admitted to Shenzhen Third People's Hospital. The study was approved by the Ethics Committee of Shenzhen Third People's Hospital (2024-195). This retrospective study was part of a prospective project, which was registered with the Chinese Clinical Trial Registry (ChiCTR, *www.chictr.org.cn*) under identifier ChiCTR2500098704.

A total of 70 suspected pulmonary TB patients from Inpatient Departments or Outpatient Clinics were enrolled in the study through the medical database within the hospital information system. Inclusion criteria were as follows: (i) inpatients or outpatients suspected to have pulmonary TB who were registered in the hospital information system and had accessible follow-up records and (ii) patients who underwent the SCMLP-TB (single cell metabolic labeling probe for tuberculosis) test using sputum or BALF samples. Eligible participants were selected based on the diagnosis criteria for pulmonary TB (WS 288-2017), with at least one of the following: symptoms consistent with TB, chest imaging compatible with TB, or a positive interferon-gamma release assay (IGRA). An experienced infectious-disease physician made this assessment. Exclusion criteria were as follows: (i) absence of a definite diagnosis after discharge or follow-up; (ii) extrapulmonary TB or severe co-infection; (iii) empirical anti-TB therapy for ≥ 2 weeks before testing; (iv) age < 18 years, pregnancy, or psychiatric disorder. Diagnostic criteria: the attending physician made a final clinical diagnosis according to the Chinese diagnostic criteria for pulmonary TB (WS 288-2017). Patients were classified into three groups: (1) confirmed TB (CT): a biological specimen was positive by any of smear microscopy, culture, Xpert MTB/RIF (Xpert), TB-DNA and metagenomic next-generation sequencing (mNGS), with non-tuberculosis mycobacteria (NTM) excluded by NTM-DNA or mNGS, (2) clinically diagnosed TB (CDxT): when patients lacked microbiological evidence, but the attending physician strongly suspected TB after excluding other diseases by combining the patient's clinical signs, chest imaging findings and other laboratory tests, patients were diagnosed with anti-TB therapy (ATT) and confirmed to have well-responded to anti-TB treatment at months of follow-up, (3) non-TB: pneumonia, malignancy, chronic obstructive pulmonary disease, bronchiectasis or other aetiologies. Pulmonary TB cases comprised CT and CDxT cases.

Data were obtained from the medical database including patients' demographics (age, gender, smoke history, comorbidities and previous TB history), clinical symptoms (cough, fever, night sweats, loss of weight or haemoptysis), chest imaging findings, laboratory results (MTB culture, Xpert, acid-fast bacilli (AFB), TB-DNA and IGRA).

2.2. Detection procedures of SCMLP-TB

Sputum and BALF samples were collected for detection. Sputum samples, treated with NaOH-NALC (DZ0802, Leagene, Beijing, China) for 15-20 minutes, aimed to reduce viscosity, remove proteins and cellular material surrounding MTB, and eliminate other bacteria, while preserving MTB due to its dense cell wall. The samples were then filtered through a 40 μm filter (SCS402, Smtra, Hangzhou, China), centrifuged at 4,000 rpm for 15 minutes, and resuspended in 100 μL phosphate buffered saline (PBS; Sangon Biotech, Shanghai, China). Next, 100 μL of the sample was mixed with 40 μM SCMLP-TB and incubated at 37°C for 2 hours to label MTB. BALF samples followed the same processing steps post-digestion, omitting the digestion stage (*22*). After labeling, samples were centrifuged at 12,000 rpm for 5 minutes and re-suspended in PBS. The labeled MTB was visualized using confocal laser scanning microscopy (LSM 980, Carl Zeiss, Oberkochen, Germany), and the fluorescence (FL) intensities of SCMLP-TB anchored on MTB were measured at 488 nm using a multi-mode microplate reader (SH1MF-SN, Shenzhen Boxing Biotechnology Co., Ltd., Shenzhen, China) (*22*).

2.3. MTB culture

At Shenzhen Third People's Hospital, BALF and sputum samples were processed using the BACTEC MGIT 960 system (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) for MTB culture. All culture tubes were cultured in a 37°C incubator. The MGIT fluorescent automatic light reader was used for interpretation. All negative culture tubes continued to be cultured and read for 8 weeks. MGIT negative culture tubes should be discarded before observing the turbidity or small particles in the tube. All positive cultures detected were confirmed by acid-fast staining. Time to positivity (TTP): the interval until the first positive signal in culture (*23*).

2.4. Xpert MTB/RIF assay

The Xpert MTB/RIF assay (Cepheid, Sunnyvale, CA, USA) was performed according to the manufacturer's instructions. One millilitre of sputum or BALF was subjected to centrifugal precipitation, and the resulting pellet was transferred to a leak-proof container and combined with 2 mL of processing solution. The mixture was vortexed vigorously for 10-20 times, then incubated at room temperature for 15 minutes. Exactly 2 mL was withdrawn with a pipette and loaded into the sample port of the reaction cassette. The lid was sealed firmly to initiate the assay. Approximately 2 hours after assay initiation, results were interpreted (*23*).

2.5. Statistical analysis

Baseline data were presented as frequency (%), mean ± SD, or median (IQR). Continuous variables were compared using independent *t*-tests for normally distributed data or Mann-Whitney *U* tests for non-normally distributed data. Categorical variables were analyzed using chi-square tests or Fisher's exact tests. To determine the fluorescence threshold for the SCMLP-TB assay, a receiver operating characteristic (ROC) curve was generated using the fluorescence intensity values from all 40 TB cases and 30 non-TB cases to determine the optimal diagnostic threshold. The optimal threshold was selected by balancing sensitivity and specificity, yielding a cutoff of 4,787 a.u. (Supplementary Figure S1, *https://www.ddtjournal.com/supplementaldata/294*). Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were reported with two-sided 95% confidence intervals (CI). Sensitivity and specificity comparisons were performed using Fisher's exact test. A *p*-value of less than 0.05 was considered to be statistically significant (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

3. Results

3.1. Study participants population

Between May 31, 2023 and October 31, 2024, 73 individuals were screened for tuberculosis (TB). The enrollment process was summarized in Figure 1. At enrollment, all participants were systematically screened for clinical symptoms and previous TB history. Of the 73 individuals, 3 were excluded: 1 with extrapulmonary TB and 2 without definitive diagnosis. Therefore, 70 individuals fulfilled the inclusion criteria and were recruited as study participants. Consequently, clinical and laboratory data were collected from these 70 screened participants at baseline (Table 1). Overall, the mean age of the study participants was 50.0 ± 16.3 years. Of the 70 participants, 29 (41.4%) were female, 41 (58.6%) were male, and 13 (18.6%) had a smoke history. Among them,

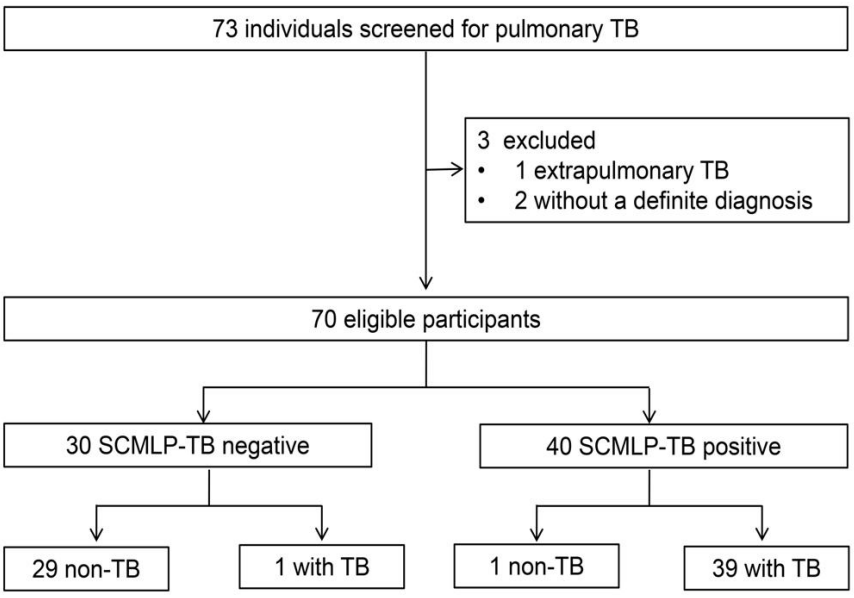

Figure 1. Flow diagram of participant enrollment. SCMLP-TB, single cell metabolic labeling probe for tuberculosis.

Table 1. Baseline characteristics of the study participants

	Number of participants with results	Overall (n = 70)	TB (n = 40)	Non-TB (n = 30)	P-value
Demographics					
Age, years	70	50.0 ± 16.3	46.5 ± 14.5	54.6 ± 17.7	0.039*
Sex	70				0.780
Female		29 (41.4%)	16 (40.0%)	13 (43.3%)	
Male		41 (58.6%)	24 (60.0%)	17 (56.7%)	
Smoke history	70	13 (18.6%)	5 (12.5%)	8 (26.7%)	0.123
Clinical					
Symptom					
Cough	70	57 (81.4%)	32 (80.0%)	25 (83.3%)	0.723
Fever	70	23 (32.9%)	11 (27.5%)	12 (40.0%)	0.264
Night sweats	70	11 (15.7%)	5 (12.5%)	6 (20.0%)	0.510
Loss of weight	70	17 (24.3%)	10 (25.0%)	7 (23.3%)	0.876
Haemoptysis	70	10 (15.7%)	6 (15.0%))	4 (16.7%)	1.000
Previous TB history	70	9 (12.9%)	6 (15.0%)	3 (10.0%)	0.723
Comorbidities					
HIV	70	13 (18.6%)	5 (12.5%)	8 (26.7%)	0.123
Diabetes	70	11 (15.7%)	10 (25.0%)	1 (3.3%)	0.011*
Haemoglobin, g/dL	70	12.2 ± 2.5	12.2 ± 2.5	12.3 ± 2.5	0.860
Positive chest imaging	68	45 (66.2%)	34 (89.5%)	11 (37.9%)	< 0.001***

Data were expressed as number (%) for categorical variables and as mean (SD) for continuous variables in case of normal distributions and median (IQR) otherwise. Positive chest imaging meant that the result of chest computed tomography was suggestive of tuberculosis. p-value: TB cases vs non-TB cases, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. TB cases were comprised of CT and CDxT. CT, confirmed tuberculosis; CDxT, clinically diagnosed tuberculosis. TB, tuberculosis.

the most frequent symptoms were cough in 57 (81.4%) and fever in 23 (32.9%), with other symptoms including night sweats in 11 (15.7%), loss of weight in 17 (24.3%), and hemoptysis in 10 (15.7%). Additionally, 9 (12.9%) of the participants had a previous history of TB prior to their enrollment in the study. Comorbidities were recorded, with HIV presenting in 13 participants (18.6%) and diabetes in 11 participants (15.7%). The mean hemoglobin level was 12.2 g/dL for the TB group and 12.3 g/dL for the non-TB group. Chest imaging findings showed that 34 out of 38 (89.5%) had abnormalities suggestive of TB in the TB group, while 11 out of 29 (37.9%) exhibited such indications in the non-TB group (Table 1).

3.2. Sensitivity and specificity analysis

We compiled detailed clinical and assay data for each participant in our study, categorizing them into confirmed TB (CT) cases, clinically diagnosed TB (CDxT) cases and non-TB cases according to the diagnosis criteria for pulmonary TB (WS 288-2017) (Figure 2A). CT and CDxT were categorized as TB cases. Since CDxT cases may introduce misclassification bias and overestimate sensitivity. To minimize this risk, final diagnoses were determined independently of SCMLP-TB results through expert adjudication. Exactly, 30 cases were CT and 10 cases were CDxT. The other 30 cases were diagnosed to be non-TB cases after follow-up visits (Figure 2A). The Venn diagram displayed 40 participants with TB detected by each diagnostic tests (culture, Xpert MTB/RIF (Xpert), and SCMLP-TB) and the overlap between tests (Figure 2B). The substantial overlap between SCMLP-TB and the other two methods indicated a strong concordance, thereby validating the reliability of SCMLP-TB alongside established diagnostic methods. Notably, the SCMLP-TB circle extended beyond the intersections, demonstrating its capacity to detect additional cases that were not identified by Xpert or culture alone. This suggested that SCMLP-TB may provide enhanced sensitivity in TB detection, complementing existing clinical tools (Figure 2B). We conducted a quantitative analysis comparing SCMLP-TB fluorescence (FL) intensity tested by microplate reader in the TB and non-TB groups (Figure 2C). 39 (97.5%) of 40 were above the detection threshold in TB group, while 1 (3.3%) of 30 in the non-TB group (Figure 2C). The median FL intensity of the TB group was 15,587 a.u., while that of the non-TB group was 2,468 a.u.. The TB group showed significantly higher FL intensity ($p < 0.001$), demonstrating robust diagnostic evidence for TB. Besides, we assessed the correlation between SCMLP-TB FL intensity and bacillary load, as measured by time to positivity (TTP). The SCMLP-TB's enhanced FL intensity showed a significant negative correlation with the TTP of culture test (Pearson's r = -0.60, p = 0.005), indicating that SCMLP-TB's FL intensity could likely serve as a reliable indicator of bacillary load in samples (Figure 2D).

Indeed, SCMLP-TB surpassed culture and Xpert assay in diagnostic performance. SCMLP-TB's sensitivity was 97.5% (95% CI 85.3-99.9; 39 of 40 cases), significantly higher than culture's 62.5% (95% CI 45.8-76.8; 25 of 40 cases) ($p < 0.05$). SCMLP-TB's specificity of 96.7% (95% CI 80.9-99.8; 29 of 30 cases) was close to culture's specificity of 100.0% (95% CI 82.8-100.0; 24 of 24 cases). Notably, SCMLP-TB detected all 10

CDxT cases missed by both culture and Xpert (Figures 2A and 2B and Table 2). This capability highlighted the potential of SCMLP-TB to serve as an ultra-sensitive diagnostic tool for paucibacillary pulmonary TB. While Xpert showed a specificity of 100% (95% CI 83.4-100; 25 of 25 cases), its sensitivity was lower at 72.5% (95% CI 55.9-84.9; 29 of 40 cases). Statistical analysis revealed that SCMLP-TB was more sensitive than Xpert ($p < 0.05$). SCMLP-TB also exhibited the optimal positive predictive value (PPV) and negative predictive value (NPV) among the three diagnostic methods, with a PPV of 97.5% (95% CI 85.3-99.9; 39 of 40 cases) and an NPV of 96.7% (95% CI 80.9-99.8; 29 of 30 cases). Culture exhibited a PPV of 100.0% (95% CI 83.4-100.0; 25 of 25 cases) and an NPV of 61.5% (95% CI 44.7-76.2; 24 of 39 cases), while Xpert showed a PPV of 100% (95% CI 85.4-100.0; 29 of 29 cases) and an NPV of 69.4% (95% CI 51.7-83.1; 24 of 35 cases), reflecting SCMLP-TB's superior diagnostic accuracy in reflecting the true TB disease status.

4. Discussion

In 2023, 10.8 million new TB cases and 1.25 million TB-related deaths occurred globally (*1*). The persistently high global incidence and mortality rates of TB highlighted the urgent need for effective diagnostic and treatment strategies. In clinical practice, Xpert MTB/RIF (Xpert),

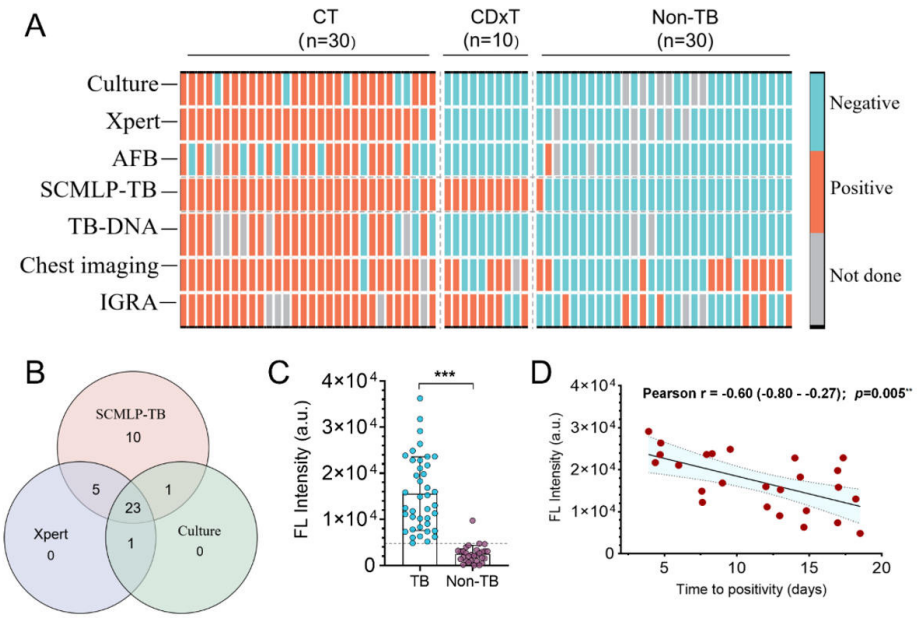

Figure 2. SCMLP-TB's diagnostic performance in participants. (**A**) Clinical and assay data for adults classified with CT cases, CDxT cases and non-TB cases. TB cases were comprised of CT and CDxT cases. (**B**) Venn diagram of overlap in TB. The Venn diagram displayed 40 participants with TB detected by each diagnostic test and the overlap between tests. (**C**) SCMLP-TB signals of samples from TB and non-TB patients. Two-sided *p* values between these two groups were obtained by the Mann-Whitney *U* tests. Dashed line denoted the threshold values ascertained by constructing ROC curve (Figure S1). (**D**) Correlation of FL intensity and the TTP of culture-positive samples (n=25). Data points represented individual patient values. Solid lines represented the linear regression line and shaded areas represented the 95% CIs. FL intensity was presented in a.u.. Pearson correlation was used to assess the statistical significance of the relationship between FL intensity and TTP. SCMLP-TB, single cell metabolic labeling probe for tuberculosis; TB, tuberculosis; CT, confirmed tuberculosis; CDxT, clinically diagnosed tuberculosis; ROC, receiver operating characteristic; FL, fluorescence; TTP, time to positivity; a.u., arbitrary units; Culture, MTB culture; Xpert, Xpert MTB/RIF; AFB, acid-fast bacilli; IGRA, interferon-gamma release assay; CIs, confidence intervals; $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Table 2. Diagnostic performance of culture, Xpert, and SCMLP-TB for tuberculosis compared to final clinical diagnosis

	Sensitivity (95% CI; n/N)	Specificity (95% CI; n/N)	PPV (95% CI; n/N)	NPV (95% CI; n/N)
Culture	62.5% (45.8-76.8; 25/40)	100.0% (82.8-100.0; 24/24)	100.0% (83.4-100.0; 25/25)	61.5% (44.7-76.2; 24/39)
Xpert	72.5% (55.9-84.9; 29/40)	100.0% (83.4-100.0; 25/25)	100.0% (85.4-100.0; 29/29)	69.4% (51.7-83.1; 25/36)
SCMLP-TB	97.5% (85.3-99.9; 39/40)	96.7% (80.9-99.8; 29/30)	97.5% (85.3-99.9; 39/40)	96.7% (80.9-99.8; 29/30)

PPV, positive predictive value; NPV, negative predictive value. Xpert, Xpert MTB/RIF. SCMLP-TB, single cell metabolic labeling probe for tuberculosis.

culture, AFB, and mNGS are used for TB diagnosis. These methods exhibited varying levels of sensitivity, with Xpert demonstrating high sensitivity (about 67-89%), culture showing limited sensitivity (50-70%), AFB exhibiting lower sensitivity (30-70%), and mNGS providing moderate sensitivity (*24-27*). These methods were still limited by the requirement for a relatively high bacterial load. Fluorescence-based metabolic labeling techniques had great potential for diagnosis and therapeutic evaluation owing to their high sensitivity (*28*). However, the chemical interaction of Auramine O with MTB was constrained by incomplete dye binding, resulting in low accuracy of only about 80% (*29*). Quantitative PCR faced challenges in differentiating live from dead bacteria and was limited by accessibility and cost in resource-limited settings (*4*). Our study designed a metabolic labeling probe (SCMLP-TB) to engage in MTB cell wall synthesis *via* antigen-85 (Ag85), achieving rapid, sensitive and accurate tuberculosis diagnosis and demonstrated great promise for point-of-care testing diagnostic applications (*11)*.

This study retrospectively evaluated the diagnostic accuracy of SCMLP-TB for the rapid detection of pulmonary TB in a real-world setting for the first time. Our data demonstrated that SCMLP-TB was an effective diagnostic tool for detecting pulmonary TB in clinical settings with 97.5% sensitivity and 96.7% specificity, significantly superior to culture and Xpert. Remarkably, the diagnostic accuracy for patients with negative culture and Xpert tests was 100%, suggesting that SCMLP-TB had excellent performance in detecting samples with low bacterial loads, such as individuals co-infected with HIV (*24*).

In this study, the sensitivities for culture and Xpert respectively were 62.5% and 72.5% compared to final clinical diagnosis by the diagnosis criteria for pulmonary tuberculosis (WS 288-2017), while both exhibited specificities of 100.0%. Indeed, our research data was essentially consistent with the reported results (*24*). Patients with subclinical TB were prone to being overlooked during screening due to the absence of obvious symptoms (*30*), which heightened the urgency to develop novel detection methods capable of identifying such patients. Previous data indicated that subclinical tuberculosis represented a median of 50.4% of all TB cases, with approximately one-third of cases in China classified as subclinical (*31,32*). Employing a combination of chest imaging and SCMLP-TB technology for asymptomatic screening is expected to enhance the detection rate of subclinical TB and is of great significance in controlling the spread of the disease. In addition, individuals co-infected with HIV are more susceptible to a variety of infections owing to their impaired immune systems, with TB being one of the most frequent opportunistic infections (*33*). A meta-analysis revealed that the sensitivity of Xpert was relatively inferior in patients with HIV-associated tuberculosis (*24*). In this study, among a total of 5 TB-HIV co-infected participants, Xpert managed to correctly diagnose merely three cases; in contrast, the SCMLP-TB test presented positive results for all five patients. Although this subgroup was too small to estimate assay performance, SCMLP-TB demonstrated a potential advantage in diagnosing TB co-infected with HIV.

While the current study illustrated the potential of SCMLP-TB in accurate TB diagnosis, it was a preliminary investigation with several limitations and thus future large-scale, multicenter prospective studies are warranted for further validation. Firstly, our study had an insufficient sample size and all the samples were from a single center. Secondly, the study failed to take children into account, thereby overlooking a substantial clinical necessity for enhanced TB diagnostic methods within pediatric populations. Thirdly, the sample size of immunosuppressed patients was limited, including only 13 individuals co-infected HIV infection and 11 with diabetes. The scarcity of participants with these specific conditions made it difficult to draw definitive conclusions about the effectiveness of the test in these sub-populations. Despite the small sample size in these subgroups, SCMLP-TB managed to diagnose all TB patients co-infected with HIV.

In conclusion, we were the first to report the application of the SCMLP-TB test on sputum or BALF specimens and showed its excellent diagnostic accuracy for pulmonary TB, particularly in patients with paucibacillary TB. *Via* one-step labeling and centrifugation, SCMLP-TB realized simple and high-throughput MTB detection by microplate reader, representing a clinically accessible and low-cost assay (~10 cents per test) for TB diagnosis. Future development will focus on an automated sample processing, fluorescence imaging, and AI-assisted analysis for efficient sample analysis.

Considering the current burden of TB and the shortcomings of existing diagnostic technologies, the implementation of the user-friendly SCMLP-TB was likely to enhance the diagnosis and treatment monitoring of TB infections that presented with a variety of symptoms and low bacterial loads.

Acknowledgements

We would like to thank all participants for their valuable contributions of biological specimens to this study.

Funding: This work was financially supported by the National Key R&D Program of China (2023YFC2308300, 2023YFA0915600), Shenzhen Medical Research Fund (D250402006, A2503081), Natural Science Foundation of China (82372271), Major Project of Guangzhou National Laboratory (GZNL2024A01009), Projects of International Cooperation and Exchanges NSFC (82561128247), Key Area Projects for Universities

in Guangdong Province (2022DZX2022), China Postdoctoral Fund (2023M731525), Guangdong Province Medical Science and Technology Research Fund (A2024400), Shenzhen Science and Technology Program (JCYJ20240813102021028, JCYJ20240813102012017, JCYJ20250604143832042) and Shenzhen Clinical Medical Center for Emerging infectious diseases (LCYSSQ20220823091203007).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. World Health Organization. Global tuberculosis report 2024. *https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024* (accessed January 21, 2026).
2. Goletti D, Meintjes G, Andrade BB, Zumla A, Shan Lee S. Insights from the 2024 WHO global tuberculosis report-more comprehensive action, innovation, and investments required for achieving WHO end TB goals. Int J Infect Dis. 2025; 150:107325.
3. Kim M, Johnson CE, Schmalstig AA, Annis A, Wessel SE, Van Horn B, Schauer A, Exner AA, Stout JE, Wahl A, Braunstein M, Victor Garcia J, Kovarova M. A long-acting formulation of rifabutin is effective for prevention and treatment of *Mycobacterium tuberculosis*. Nat Commun. 2022; 13:4455.
4. Musisi E, Wamutu S, Ssengooba W, *et al*. Accuracy of the tuberculosis molecular bacterial load assay to diagnose and monitor response to anti-tuberculosis therapy: a longitudinal comparative study with standard-of-care smear microscopy, Xpert MTB/RIF Ultra, and culture in Uganda. Lancet Microbe. 2024; 5:e345-e354.
5. Shah M, Paradis S, Betz J, *et al*. Multicenter study of the accuracy of the BD MAX multidrug-resistant tuberculosis assay for detection of *Mycobacterium tuberculosis* complex and mutations associated with resistance to rifampin and isoniazid. Clin Infect Dis. 2020; 71:1161-1167.
6. Vongthilath-Moeung R, Poncet A, Renzi G, Schrenzel J, Janssens JP. Time to detection of growth for *Mycobacterium tuberculosis* in a low incidence area. Front Cell Infect Microbiol. 2021; 11:704169.
7. Youngquist BM, Saliba J, Kim Y, *et al*. Rapid tuberculosis diagnosis from respiratory or blood samples by a low cost, portable lab-in-tube assay. Sci Transl Med. 2025; 17:eadp6411.
8. Denkinger CM, Schumacher SG, Boehme CC, Dendukuri N, Pai M, Steingart KR. Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis. Eur Respir J. 2014; 44:435-446.
9. Kohli M, Schiller I, Dendukuri N, Dheda K, Denkinger CM, Schumacher SG, Steingart KR. Xpert® MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. Cochrane Database Syst Rev. 2018; 8:CD012768.
10. Lawn SD, Brooks SV, Kranzer K, Nicol MP, Whitelaw A, Vogt M, Bekker LG, Wood R. Screening for HIV-associated tuberculosis and rifampicin resistance before antiretroviral therapy using the Xpert MTB/RIF assay: a prospective study. PLoS Med. 2011; 8:e1001067.
11. Banahene N, Gepford DM, Biegas KJ, Swanson DH, Hsu YP, Murphy BA, Taylor ZE, Lepori I, Siegrist MS, Obregón-Henao A, Van Nieuwenhze MS, Swarts BM. A far-red molecular rotor fluorogenic trehalose probe for live mycobacteria detection and drug-susceptibility testing. Angew Chem Int Ed Engl. 2023; 62:e202213563.
12. Barr DA, Schutz C, Balfour A, *et al*. Serial measurement of *M. tuberculosis* in blood from critically-ill patients with HIV-associated tuberculosis. EBioMedicine. 2022; 78:103949.
13. Kamariza M, Shieh P, Ealand CS, Peters JS, Chu B, Rodriguez-Rivera FP, Babu Sait MR, Treuren WV, Martinson N, Kalscheuer R, Kana BD, Bertozzi CR. Rapid detection of *Mycobacterium tuberculosis* in sputum with a solvatochromic trehalose probe. Sci Transl Med. 2018; 10:eaam6310.
14. Welsh KJ, Hunter RL, Actor JK. Trehalose 6,6'-dimycolate--a coat to regulate tuberculosis immunopathogenesis. Tuberculosis. 2013; 93:S3-S9.
15. Rodriguez-Rivera FP, Zhou X, Theriot JA, Bertozzi CR. Visualization of mycobacterial membrane dynamics in live cells. J Am Chem Soc. 2017; 139:3488-3495.
16. Kamariza M, Keyser SGL, Utz A, Knapp BD, Ealand C, Ahn G, Cambier CJ, Chen T, Kana B, Huang KC, Bertozzi CR. Toward point-of-care detection of *Mycobacterium tuberculosis*: a brighter solvatochromic probe detects mycobacteria within minutes. JACS Au. 2021; 1:1368-1379.
17. Belisle JT, Vissa VD, Sievert T, Takayama K, Brennan PJ, Besra GS. Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. Science. 1997; 276:1420-1422.
18. Oliveira E, Bértolo E, Núñez C, Pilla V, Santos HM, Fernández-Lodeiro J, Fernández-Lodeiro A, Djafari J, Capelo JL, Lodeiro C. Green and red fluorescent dyes for translational applications in imaging and sensing analytes: a dual-color flag. ChemistryOpen. 2017; 7:9-52.
19. Zhao N, Lam JW, Sung HH, Su HM, Williams ID, Wong KS, Tang BZ. Effect of the counterion on light emission: a displacement strategy to change the emission behaviour from aggregation-caused quenching to aggregation-induced emission and to construct sensitive fluorescent sensors for Hg^{2+} detection. Chemistry. 2014; 20:133-138.
20. Huang Y, Ma X, Gao C, Xie Y, Qiu M, Yin J. Fluorescent probes for glucolipid metabolism of bacterial cell wall. Green Chem Eng. 2023; 4:417-426.
21. Dai G, Luo Y, Liao M, *et al*. A cytoderm metabolic labeling AIEgen for rapid detection and intracellular ablation of *Mycobacterium tuberculosis*. Cell Rep Phys Sci. 2023; 4:101238.
22. Yang M, Dai G, Li D, Zhao P, Zhan S, Qin H, Lu H, Zheng M, Zhang P. A cytoderm metabolic labeling TPAPy-Tre for real-time detection of vitality of *Mycobacterium tuberculosis* in sputum. Microbiol Spectr. 2025; 13:e0245724.
23. Shi CL, Han P, Tang PJ, Chen MM, Ye ZJ, Wu MY, Shen J, Wu HY, Tan ZQ, Yu X, Rao GH, Zhang JP. Clinical metagenomic sequencing for diagnosis of pulmonary tuberculosis. J Infect. 2020; 81:567-574.
24. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev. 2014; 2014:CD009593.
25. Sun Y, Zhang Q, Zhang Q, Liu C, Zhang H, Fu Y, Liu Y, Hou G. Diagnostic efficacy of Xpert MTB/RIF assay in bronchoalveolar lavage fluid for tracheobronchial

tuberculosis: a retrospective analysis. Front Med. 2021; 8:682107.
26. Steingart KR, Ng V, Henry M, Hopewell PC, Ramsay A, Cunningham J, Urbanczik R, Perkins MD, Aziz MA, Pai M. Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis. 2006; 6:664-674.
27. Li Y, Bian W, Wu S, Zhang J, Li D. Metagenomic next-generation sequencing for *Mycobacterium tuberculosis* complex detection: a meta-analysis. Front Public Health. 2023; 11:1224993.
28. Liu H, Xiong LH, Kwok RTK, He X, Lam JWY, Tang BZ. AIE bioconjugates for biomedical applications. Adv Opt Mater. 2020; 8:2000162.
29. Marais BJ, Brittle W, Painczyk K, Hesseling AC, Beyers N, Wasserman E, van Soolingen D, Warren RM. Use of light-emitting diode fluorescence microscopy to detect acid-fast bacilli in sputum. Clin Infect Dis. 2008; 47:203-207.
30. Appleton SC, Connell DW, Singanayagam A, Bradley P, Pan D, Sanderson F, Cleaver B, Rahman A, Kon OM. Evaluation of prediagnosis emergency department presentations in patients with active tuberculosis: the role of chest radiography, risk factors and symptoms. BMJ Open Respir Res. 2017; 4:e000154.
31. Frascella B, Richards AS, Sossen B, Emery JC, Odone A, Law I, Onozaki I, Esmail H, Houben RMGJ. Subclinical tuberculosis disease-a review and analysis of prevalence surveys to inform definitions, burden, associations, and screening methodology. Clin Infect Dis. 2021; 73:e830-e841.
32. Tang P, Liang E, Zhang X, Feng Y, Song H, Xu J, Wu M, Pang Y. Prevalence and risk factors of subclinical tuberculosis in a low-incidence setting in China. Front Microbiol. 2022; 12:731532.
33. Perez F, Caceres DH, Ford N, Ravasi G, Gomez BL, Pasqualotto AC, Hine P, Adenis AA, Nacher M, Chiller T, Baddley J, for the Guideline Development Group for diagnosing and managing disseminated histoplasmosis among people living with HIV. Summary of guidelines for managing histoplasmosis among people living with HIV. J Fungi (Basel). 2021; 7:134.

Received November 15, 2025; Revised February 16, 2026; Accepted April 2, 2026.

[§]These authors contributed equally to this work.

**Address correspondence to:*
Hongzhou Lu and Pengfei Zhao, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital, Southern University of Science and Technology, Shenzhen, 518112, China.
E-mail: luhongzhou@fudan.edu.cn (HL); pengfeizhao@aliyun.com (PZ)

Yang Zhou, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital, Southern University of Science and Technology, Shenzhen, 518112, China; Molecular Biology Research Center & Center for Medical Genetics, School of Life Sciences, Central South University, Changsha, 410078, China.
E-mail: yangzhou@szsy.sustech.edu.cn

Mingbin Zheng, National Clinical Research Center for Infectious Disease, Shenzhen Third People's Hospital, Southern University of Science and Technology, Shenzhen, 518112, China; The Affiliated Dongguan Songshan Lake Central Hospital, Guangdong Medical University, Dongguan, 523808, China.
E-mail: mingbinzheng@126.com

Released online in J-STAGE as advance publication April 7, 2026.

Original Article

DOI: 10.5582/ddt.2026.01007

Chemotherapy combined targeted therapy has a potential benefit on the survival of advanced intrahepatic cholangiocarcinoma

Jian Yang[1,2,§], Tengqian Tang[3,§], Rui Liao[2,§], Yan He[4], Hua Zhang[2,*]

[1] Department of Hepatobiliary Surgery, Nanchuan Hospital of Chongqing Medical University, Chongqing, China;
[2] Department of Hepatobiliary Surgery, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China;
[3] Department of Hepatobiliary Surgery, the First Affiliated Hospital of Army Medical University, Chongqing, China;
[4] Sichuan College of Traditional Chinese Medicine, Mianyang, China.

SUMMARY: The treatment strategies and outcomes for advanced intrahepatic cholangiocarcinoma (ICC) are restricted. The treatment modalities and effectiveness of local interventional chemotherapy and systemic chemotherapy remain indeterminate. The objective of this study is to explore and assess the influence of different chemotherapy methods on the prognosis of patients with advanced ICC. A retrospective investigation was carried out at two research centers. The recruited patients were divided into the systemic chemotherapy cohort and the local chemotherapy cohort (Transarterial Chemoembolization, TACE). The primary endpoint of this study was overall survival (OS), while the secondary endpoints encompassed progression free survival (PFS), response rate (RR), and adverse events (AE). From January 2014 to January 2024, a total of 124 patients were included. Systemic/local chemotherapy combined with targeted therapy exhibited superior survival performance compared to chemotherapy alone. Additionally, patients with lesions confined to the liver and large tumors (> 6 cm) obtained better survival benefits from systemic chemotherapy. There was no significant disparity in grade 3 or more severe adverse events between the two groups. Whether it is systemic chemotherapy or TACE, combining them with targeted therapy can confer significant therapeutic advantages to patients with advanced ICC. ICC patients with a higher tumor burden may attain better therapeutic outcomes by selecting systemic chemotherapy.

Keywords: intrahepatic cholangiocarcinoma, liver malignant tumor, transarterial chemoembolization, local chemotherapy, systemic chemotherapy, targeted therapy, immunotherapy

1. Introduction

Intrahepatic cholangiocarcinoma (ICC), originating from the secondary branches of the bile duct, stands as the second most prevalent malignant tumor of the biliary tract, accounting for 10%-15% of primary liver cancers (*1*). In recent years, a notable upsurge in the incidence of ICC has been observed globally (*2*). Surgical resection represents the sole therapeutic approach with the potential to cure ICC patients. Nevertheless, in China, merely 20% of patients present indications for radical resection at the time of initial diagnosis (*3*). Among patients undergoing radical resection, the median survival period is 27 months, and the 5-year survival rate stands at 31% (*4*). Conversely, patients undergoing systemic chemotherapy exhibit a median overall survival of less than 1 year, and in the absence of treatment, the median survival is less than 5 months (*5,6*).

For patients with locally advanced or metastatic ICC, there exists no unified standard treatment protocol. Historically, systemic treatment for advanced biliary tract cancer (BTC) was confined to chemotherapy. A phase III clinical study from the United Kingdom, namely ABC-02, compared gemcitabine plus cisplatin with gemcitabine monotherapy in the treatment of locally advanced or metastatic BTC. Gemcitabine combined with cisplatin (GC) demonstrated a significant survival advantage and was established as the standard treatment regimen for BTC in first-line systemic therapy (*5*). Notably, a multicenter comparative study in Japan also verified the efficacy and safety of the GC treatment regimen in Japanese BTC patients, and it was adopted as the standard treatment regimen for BTC (*7*). Subsequently, in a randomized phase III study carried out in Japan to explore diverse first-line treatment options for BTC, the GC treatment regimen was compared with gemcitabine combined with tegafur-uracil (GS) chemotherapy. The median overall survival (OS) was 13.4 months and 15.1 months ($p = 0.046$), respectively, and GS was adopted as the standard treatment for patients with advanced BTC

(*8*). Moreover, a Japanese phase III study confirmed that the triple therapy of gemcitabine combined with cisplatin and teysuno (S-1) was superior to GC and was adopted as the standard treatment regimen for patients with advanced BTC (*9*). A post hoc analysis of a prospective, randomized clinical study on advanced biliary cancer named ABC-01, ABC-02, and ABC-03 revealed that patients with advanced ICC had better OS compared to other patients with extrahepatic biliary cancer (*5,10,11*). However, no large-scale prospective trials comparing local chemotherapy (TACE) with systemic chemotherapy have been conducted for patients with advanced ICC.

In recent years, the rapid advancement of immunotherapy and targeted therapy has substantially altered the overall scenario of BTC treatment. In a phase III study, namely TOPAZ-1, it was proven that durvalumab in combination with GC provided sustained and stable OS advantages. This treatment regimen marked the entry of systemic treatment for BTC into the immunotherapy era and was the first and sole global phase III clinical study to attain comprehensive survival benefits in terms of OS and progression-free survival (PFS) (*12*). In another global phase III study, KEYNOTE-966, it was revealed that, in comparison with GC monotherapy, pembrolizumab combined with GC significantly enhanced the OS of patients, and the difference was statistically significant. This discovery implies that this treatment regimen is anticipated to become a novel treatment alternative for previously untreated patients with metastatic or unresectable BTC (*13*).

In the present study, the objective was to compare the efficacy and survival advantages of patients with unresectable ICC who underwent local chemotherapy or systemic chemotherapy. The aim is to determine the optimal treatment regimen for patients with advanced ICC under various circumstances.

2. Materials and Methods

2.1. Study design and patients

A retrospective analysis was carried out on the prognosis of 124 patients with ICC who were diagnosed and TACE or systemic chemotherapy at the First Affiliated Hospital of Chongqing Medical University and the First Affiliated Hospital of Army Medical University from January 2014 to January 2024. The study protocol was approved by the Clinical Research Ethics Committee of Chongqing Medical University (K2023-451-02) and followed the Declaration of Helsinki. Informed consents were obtained from all patients.

All patients were newly diagnosed ICC cases without any other concomitant tumors and had not received any prior anti - tumor treatment. Additionally, all patients were aged between 18 and 80 years and had an Eastern Cooperative Oncology Group (ECOG) performance status ranging from 0 to 2. Patients with a history of prior chemotherapy or radiotherapy were excluded. Those with severe pulmonary or cardiac diseases were also excluded. Pregnant or lactating women were excluded too. Baseline characteristic data of the research subjects were collected, including patients' survival status, age, gender, tumor location and size, lymph node invasion status, vascular invasion status, metastasis sites, alterations in serological indicators, follow-up imaging data, and the selection of different chemotherapy regimens (Table 1).

2.2. Procedures

Regarding the treatment options for patients with inoperable tumors in this study, each center made decisions through multidisciplinary consultation and deliberation. The treatment modalities included systemic chemotherapy or local chemotherapy. Systemic chemotherapy regimens primarily involved intravenous infusion of chemotherapy drugs (single or combined treatment based on gemcitabine, oxaliplatin, tegafur), combined with targeted therapy and immunotherapy. TACE entailed first injecting chemotherapy drugs and iodized oil, followed by using drug - eluting beads for embolization to maintain a high concentration of chemotherapy drugs in the tumor area until complete stasis of the tumor's blood supply vessels was attained. Subsequently, targeted therapy and immunotherapy were combined. The patients in this study were initially staged, and the evidence of extra-hepatic lesions was verified through enhanced CT and enhanced MRI imaging. Then,

Table 1. Baseline characteristics of the patients' clinical pathology

Characteristics	*n* = 124	Percentage (%)
Gender		
Male	71	57.3
Female	53	42.7
Median age (range)	58	(35-80)
Treatment method		
Systemic treatment	88	71.0
TACE	36	29.0
Tumor characteristics		
Median size (range)	7.0	2.3-15.4
Multiple lesions	100	80.6
Long-distance transfer	59	47.6
Lymph node	95	76.6
Vascular invasion	89	71.8
Clinical picture		
Abdominal pain	54	43.5
Liver cirrhosis	66	53.2
Jaundice	21	16.9
Hepatitis B	44	35.5
CA19-9	84	67.7
AFP	14	11.3
Imaging		
CT	124	100
MRI	124	100

Abbreviations: TACE: Transarterial Chemoembolization; CA19-9: Carbohydrate Antigen19-9; AFP: Alpha-Fetal Protein; CT: Computed Tomography; MRI: Magnetic Resonance Imaging.

a biopsy was performed for pathological confirmation.

2.3. Follow-up

The median follow-up time was 11 months. After the treatment commenced, the changes in the lesions were monitored through imaging (enhanced CT and enhanced MRI) every 6-8 weeks. For patients with measurable target lesions, the tumor was evaluated in accordance with the Response Evaluation Criteria in Solid Tumors (RECIST) 1.1. When the disease progressed, the treatment plan was adjusted according to specific circumstances.

2.4. Outcomes

In treatment modalities entailing the continuous administration of local chemotherapy or systemic chemotherapy, continuous surveillance of the patient's condition is conducted until the following scenarios transpire: progression of intrahepatic lesions, extrahepatic progression, severe treatment-associated toxicity and complications, successful conversion followed by surgical resection, or patient demise. The primary endpoint of this research is overall survival (OS), while the secondary endpoints are progress free survival (PFS) and adverse events (AEs). OS is delineated as the duration commencing from the date of diagnosis until death due to any cause. PFS is defined as the period from the date of diagnosis until tumor progression or death from any cause.

2.5. Statistical analysis

Survival analyses were carried out for patients in the local chemotherapy group and the systemic chemotherapy group according to different treatment regimens. SPSS version 26.0 was used for statistical analysis. Categorical data were expressed as numbers (%), and group comparisons were performed using x^2 test or the Fisher exact probability method. Normally distributed measurement data was evaluated using the Kolmogorov-Smirnov test. Normally distributed measurement data were presented as mean ± SD (x ± s) and analyzed using the *t* test for two independent samples. Non-normally distributed measurement data were presented as the median (Q1, Q3) and compared using the rank sum test. Survival curves were plotted *via* the Kaplan-Meier method, and the statistical disparities in survival outcomes were computed. The log-rank test was employed to assess the differences between survival curves. $p < 0.05$ was statistically significant.

3. Results

3.1. Baseline characteristics

Following the exclusion of patients who did not receive any treatment (n = 53) and those who solely received radiofrequency ablation treatment (n = 9), a total of 124 patients were ultimately incorporated into the survival analysis (Figure 1). In Table 1, the median age of the included patients was 58 years (35-80 years). There were 71 males (57.3%) and 53 females (42.7%). The median size of the tumors detected *via* imaging was 7.0 cm (2.3-15.4 cm). According to the Eastern Cooperative Oncology Group Performance Status (ECOG), all patients had scores ranging from 0 to 2. The most prevalent clinical manifestations were as follows: abdominal pain was observed in 54 patients (43.5%); multiple tumor foci were present in 100 patients (80.6%); extrahepatic distant metastasis was detected in 59 patients (47.6%); vascular invasion was identified in 89 patients (71.8%); lymph node invasion was found in 95 patients (76.6%); liver cirrhosis was diagnosed in 66 patients (53.2%); and hepatitis B was confirmed in 44 patients (35.5%).

3.2. Chemotherapy regimen

Among patients undergoing total body chemotherapy, 45 patients were administered a first line chemotherapy regimen centered around gemcitabine. Specifically, 21 patients (47%) received a combination of gemcitabine and cisplatin, 9 patients (20%) received a combination of gemcitabine and albumin-bound paclitaxel, 9 patients (20%) received a combination of gemcitabine and oxaliplatin, 5 patients (11%) received a combination of gemcitabine and tegafur-uracil, and 1 patient (2%) received gemcitabine monotherapy. In addition, 15 patients were treated with a chemotherapy regimen based on oxaliplatin. Twelve patients were subjected to a systemic treatment regimen based on S-1, and 22 patients received other systemic chemotherapy regimens (Table 2).

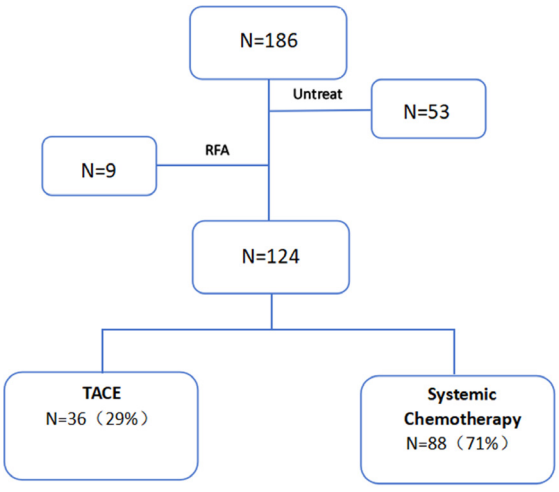

Figure 1. Grouping according to different chemotherapy modalities.

We further classified the diverse chemotherapy regimens into subgroups for survival comparison. Through comparative analysis, it was found when comparing gemcitabine plus albumin paclitaxel chemotherapy regimen with TACE, the median survival time was 9 months (7-17 months) as opposed to 11 months (2-66 months); $p = 0.05$. The survival advantage of TACE regimen was more prominent than that of gemcitabine plus albumin paclitaxel chemotherapy. Moreover, it was noted that the median overall survival (mOS) of the gemcitabine combined with cisplatin regimen was 12 months (2-48 months), indicating better survival outcomes compared to other regimens.

Table 2. Whole-body chemotherapy regimen

Chemotherapy regimen	*n* (%)
Gemcitabine-based	45 (100)
GP	21 (47)
AG	9 (20)
GEMOX	9 (20)
GS	5 (11)
Gemcitabine	1 (2)
Oxaliplatin-based	15 (100)
SOX	3 (20)
XELOX	3 (20)
GEMOX	9 (60)
S-1-based	12 (100)
S-1	4 (33)
SOX	3 (25)
GS	5 (42)
Other systemic drugs	22 (100)

Abbreviations: GP: Gemcitabine combined with Cisplatin; AG: Gemcitabine combined with albumin paclitaxel; GEMOX: Gemcitabine combined with Oxaliplatin; GS: Gemcitabine combined with Tarceva; SOX: Oxaliplatin combined with tegafur-oguacitinib; XELOX: Oxaliplatin combined with Capecitabine; S-1: Tegafur; Other systemic drugs: irinotecan, capecitabine, calcium folinate, *etc.*

3.3. Therapeutic effect of all patients

The clinical characteristics of the 124 ICC patients who underwent systemic chemotherapy or local chemotherapy in this study are presented in Table 3. Among them, 88 patients (71%) opted for systemic chemotherapy, whereas 36 patients (29%) selected TACE. Regarding patients receiving systemic chemotherapy, 39 (44.3%) were intolerant to the combined targeted and immunotherapy and only received systemic chemotherapy with chemotherapeutic agents. Forty nine (55.7%) chose systemic chemotherapy combined with targeted and immunotherapy, and the treatment regimens were all determined based on the results of genetic testing.

In this study, 99 (79.8%) patients succumbed to ICC. The mOS of all patients in this study was 11 months (1-66 months). The mOS and Median mPFS in the systemic chemotherapy group were 10 months (range, 2-53 months) and 6 months (range, 2-49 months), respectively. The mOS and mPFS of the TACE group were 11 months (1-66 months) and 6 months (1-57 months), respectively.

3.4. Systemic/local chemotherapy combined with targeted therapy had better survival performance than chemotherapy alone

In the survival analysis comparing patients undergoing TACE combined with targeted therapy and those receiving chemotherapy alone (Figure 2a), the mOS was 19 months (7-66 months) for the former group and 6 months (2-15 months) for the latter group, $p < 0.05$. Patients receiving TACE combined with targeted therapy exhibited superior survival performance. In the survival analysis comparing patients receiving TACE combined with targeted therapy and those undergoing systemic chemotherapy combined with targeted therapy

Table 3. Tumor characteristics and performance status score

	TACE (*n* = 36)	Systemic treatment (*n* = 88)	*p*
Median tumor size (range)	7.8 (2.8-15.0)	7.8 (2.8-15.0)	0.34
Number of tumors			< 0.05
single	19 (52.8%)	19 (52.8%)	
multiple	17 (47.2%)	17 (47.2%)	
Lymph node metastasis			< 0.05
No	16 (44.4%)	16 (44.4%)	
Local	7 (19.4)	7 (19.4)	
Multiple	13 (36.1%)	13 (36.1%)	
Extrahepatic metastasis	22 (61.1%)	22 (61.1%)	0.03
Vascular involvement	24 (66.7%)	24 (66.7%)	0.56
CA19-9	16 (44.4%)	16 (44.4%)	< 0.05
Liver cirrhosis	12 (33.3%)	12 (33.3%)	< 0.05
Median treatment cycle	2 (1-5)	2 (1-5)	0.19
Combined targeted immunotherapy	23 (63.9%)	23 (63.9%)	0.52
ECOG			0.11
0	22 (61.1%)	22 (61.1%)	
1	12 (33.3%)	12 (33.3%)	
2	2 (5.6%)	2 (5.6%)	

Abbreviations: TACE: Transarterial Chemoembolization; CA19-9: Carbohydrate Antigen19-9; ECOG: Eastern Cooperative Oncology Group Performance Status Score.

(Figure 2b), the mOS was 19 months (7-66 months) and 16 months (4-53 months) respectively, $p = 0.39$. There was no significant difference in survival between the two groups. When a survival analysis was carried out for patients receiving only TACE and those receiving systemic chemotherapy combined with targeted therapy (Figure 2c), the mOS was 5 months (1-11 months) for the former group and 16 months (4-53 months) for the latter group, $p < 0.05$. The survival benefit of systemic chemotherapy combined with targeted therapy was significantly greater than that of receiving only TACE. In the survival analysis comparing patients receiving only TACE and those receiving only chemotherapy (Figure 2d), the mOS was 5 months (1-11 months) and 6 months (2-15 months) respectively, $p = 0.68$. There was no significant difference in survival between the two groups. In the survival analysis comparing patients receiving only TACE and those receiving TACE combined with targeted therapy (Figure 2e), the mOS was 5 months (1-11 months) for the former group and 19 months (7 -

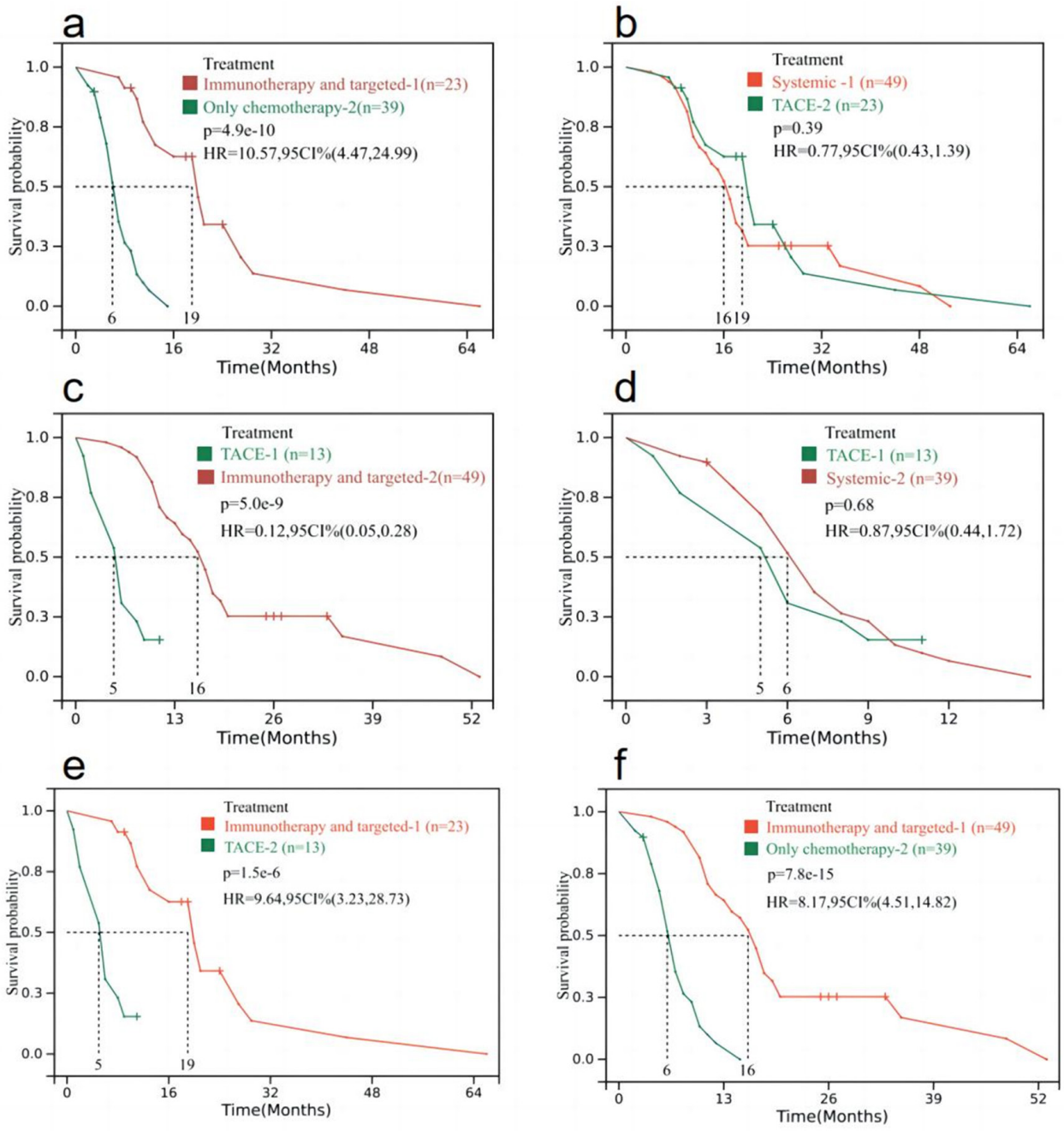

Figure 2. Survival curves of the TACE and systemic chemotherapy subgroups. a: 1 is the TACE combined with targeted immunotherapy; 2 is chemotherapy alone, $p < 0.05$; **b**: 1 is systemic chemotherapy combined with targeted immunotherapy; 2 is TACE combined with targeted immunotherapy, $p = 0.39$; **c**: 1 is the sole TACE regimen; 2 is systemic chemotherapy combined with targeted immunotherapy, $p < 0.05$; **d**: 1 is the sole TACE regimen; 2 is chemotherapy alone, $p = 0.68$; **e**: 1 is the TACE combined with targeted immunotherapy; 2 is the sole TACE regimen, $p < 0.05$; **f**: 1 is systemic chemotherapy combined with targeted immunotherapy; 2 is chemotherapy alone, $p < 0.05$.

66 months) for the latter group, $p < 0.05$. The survival benefit of TACE combined with targeted therapy was significantly greater than that of receiving only TACE. When a survival analysis was conducted for patients receiving only chemotherapy and those receiving chemotherapy combined with targeted therapy (Figure 2f), the mOS was 6 months (2-15 months) for the former group and 16 months (4-53 months) for the latter group; $p < 0.05$. The survival benefit of patients receiving systemic chemotherapy combined with targeted therapy was significantly greater than that of patients receiving only chemotherapy.

3.5. Confined lesions in the liver and large tumor (> 6 cm) have a better survival benefit from systemic chemotherapy

In the survival analysis incorporating liver and extrahepatic organ metastases of tumors, among patients undergoing systemic chemotherapy (Figure 3a), the mOS of patients with liver-only lesions was 11 months (4-35 months) compared with 8 months (2-48 months) for patients with other types of lesions. $p < 0.05$. Lesions restricted to the liver conferred a more favorable survival benefit compared to those in extrahepatic organs. When examining the tumor location in patients receiving TACE (Figure 3b), the mOS of patients with lesions restricted to the liver was 8 months (1-44 months), as opposed to 13 months (2-66 months); $p = 0.75$. Among patients receiving local chemotherapy, there was no significant disparity in survival benefit between those with lesions restricted to the liver and those with extrahepatic organ metastases.

Upon re-incorporating the maximum tumor diameter into the survival analysis, among patients who underwent systemic chemotherapy, with the maximum tumor diameter demarcated at 6 cm, it was discovered that patients with a maximum tumor diameter exceeding 6 cm exhibited more favorable survival benefits compared to those with a maximum tumor diameter less than 6 cm (Figure 4). The mOS was 12 months (2-53 months) as opposed to 8 months (2-18 months); $p < 0.05$. Nevertheless, no survival disparities were detected

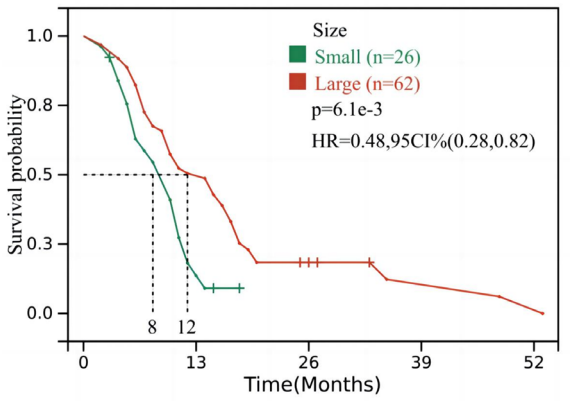

Figure 4. Survival curve of the maximum tumor diameter after systemic chemotherapy.

Table 4. Baseline characteristics of patients based on the maximum diameter of the tumor

Tumor diameter (cm)	< 6 cm (n = 26)	> 6 cm (n = 62)	p
Lymph	24 (92.3%)	51 (82.3%)	0.38
Extrahepatic	6 (23.1%)	25 (40.3%)	0.19
Vascular involvement	19 (73.1%)	46 (74.2%)	1.00
Number of lesions			0.98
1	2 (7.7%)	3 (4.8%)	
≥ 2	24 (92.3%)	59 (95.2%)	
Targeted immunotherapy	14 (53.8%)	35 (56.5%)	1.00
Median treatment cycle	4.0 (2.0-6.0)	3.0 (1.8-6.0)	0.69
ECOG			0.18
0	9 (34.6%)	31 (50.0%)	
1	9 (34.6%)	22 (35.4%)	
2	8 (30.7%)	9 (14.5%)	

Abbreviations: ECOG: Eastern Cooperative Oncology Group Performance Status Score.

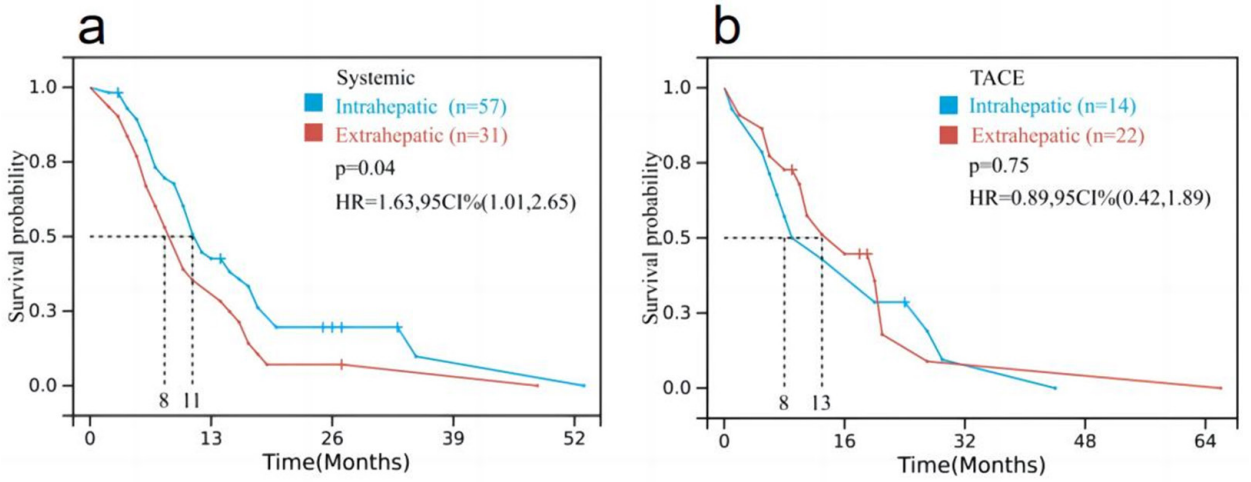

Figure 3. Survival curve of tumor metastasis sites. a: For patients undergoing systemic chemotherapy; **b**: For patients receiving TACE treatment.

among patients with varying tumor diameters in the TACE group. Table 4 show cases the subsequent analysis of the tumor characteristics of patients who received systemic chemotherapy, with a maximum tumor diameter of 6 cm serving as the standard.

3.6. Treatment-related adverse events (TRAE)

An analysis of treatment-related adverse events (TRAE) was conducted on all patients who received treatment (Table 5). Among the 124 patients included in this study, 47 (37.9%) experienced treatment-related adverse events. Among these 47 patients, 38 (43.2%) had grade 1-2 events, and 9 (10.2%) had grade 3 or above adverse events. The grade 3 adverse events primarily manifested as lymphocyte reduction in 6 (4.8%) patients, neutrophil reduction in 6 (4.8%) patients, and platelet reduction in 8 (6.5%) patients. This may be associated with the bone marrow suppression induced by chemotherapy drugs. There were 12 (9.7%) patients with immune-related adverse events (irAEs), mainly presenting as immune-related pneumonia after treatment. In addition, the majority of patients (77, 62.1%) experienced weight loss after treatment, all of which were grade 1-2 adverse events. Among the patients who received TACE, the main manifestation was liver function impairment after embolization, all of which were grade 1-2 adverse events. Alanine aminotransferase increased in 14 (38.9%) patients; aspartate aminotransferase increased in 14 (38.9%) patients. All the related adverse events were resolved after treatment discontinuation and symptomatic treatment.

4. Discussion

In the past three years, the core frontier of ICC research has centered on chemotherapy combined with targeted immunotherapy and stratified treatment based on tumor characteristics (*e.g.*, tumor size, metastasis site, and the influence of physical condition on treatment selection) (*14*). By analyzing subgroups of patients undergoing systemic chemotherapy or TACE, it was found that both local chemotherapy combined with targeted immunotherapy and systemic chemotherapy combined with targeted immunotherapy resulted in more favorable long-term survival and PFS outcomes compared to patients who received only chemotherapy drugs or only TACE. This phenomenon might be ascribed to the support of targeted combined immunotherapy. As confirmed in the "TOPAZ-1" study, durvalumab plus GC had superior OS compared with placebo, with a mOS of 12.9 months (11.6-14.1 months) *vs.* 11.3 months (10.1-12.5 months) ($p < 0.05$). In the KEYNOTE-966 study, pembrolizumab combined with GC had a better overall survival compared with placebo in patients with advanced biliary tract cancer, with a mOS of 12.7 months (11.5-13.6 months) and 10.9 months (9.9-11.6 months), respectively ($p < 0.05$) (*12,13*). In our study, patients who received systemic chemotherapy combined with targeted immunotherapy had a better survival than those who received chemotherapy alone, 16 months (4-53 months) *vs.* 6 months (2-15 months) ($p < 0.05$). Systemic chemotherapy combined with targeted immunotherapy also has better survival performance.

Although there are currently no specific targeted

Table 5. Analysis of treatment-related adverse events

Events *n* (%)	TACE *n* = 36		Systemic *n* = 88	
	Grade 1–2	Grade 3 or higher	Grade 1–2	Grade 3 or higher
Lymphocytes	5 (13.9)	0 (0)	20 (22.7)	5 (5.7)
Neutrophils	6 (16.7)	1 (2.8)	21 (23.9)	5 (5.7)
PLT	8 (22.2)	1 (2.8)	23 (26.1)	7 (8.0)
Anemia	4 (11.1)	1 (2.8)	12 (13.6)	2 (2.3)
ALB	5 (13.9)	1 (2.8)	19 (21.6)	3 (3.4)
ALT	12 (33.3)	0 (0)	10 (11.4)	0 (0)
AST	12 (33.3)	0 (0)	10 (11.4)	0 (0)
TBIL	10 (27.8)	0 (0)	9 (10.2)	0 (0)
Hypertension	10 (27.8)	0 (0)	14 (15.9)	0 (0)
High blood sugar	9 (25.0)	0 (0)	12 (13.6)	0 (0)
Anorexia	5 (13.9)	0 (0)	17 (19.3)	0 (0)
Nausea	6 (16.7)	0 (0)	10 (11.4)	0 (0)
Diarrhea	7 (19.4)	0 (0)	11 (12.5)	0 (0)
Oral ulcer	7 (19.4)	0 (0)	9 (10.2)	0 (0)
URTI	3 (8.3)	1 (2.8)	8 (9.1)	3 (3.4)
Cough	4 (11.1)	0 (0)	9 (10.2)	0 (0)
Fatigue	6 (16.7)	0 (0)	21 (23.9)	0 (0)
Weight loss	8 (22.2)	0 (0)	69 (78.4)	0 (0)
Rash	7 (19.4)	0 (0)	13 (14.8)	0 (0)
Abdominal pain	8 (22.2)	0 (0)	6 (6.8)	0 (0)

Abbreviations: TACE: Transarterial Chemoembolization; PLT: platelet; ALB: albumin; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; TBIL: Total bilirubin; URTI: upper respiratory tract infection.

drugs for ICC, up to 40% of patients with biliary system malignant tumors possess potential targetable genetic variations, amplifications, and fusions, including FGFR2 fusion/rearrangement, IDH mutation, NTRK fusion, HER2 amplification/expression, BRAF mutation, RET fusion, *etc.* (*14-17*). Through genetic testing, some patients can derive benefits from targeted drugs *via* the discovery of variant genes. However, this study did not conduct in-depth gene analysis, which hindered us from determining the treatment options for patients with different molecular subtypes. Additionally, although immunotherapy provides limited long - term survival benefits for ICC patients, the combination of chemotherapy and targeted drug treatment can enhance OS.

Within the framework of systemic chemotherapy, patients with lesions restricted to the liver demonstrate higher treatment efficacy and longer survival durations in comparison to those with extrahepatic and lymph node metastases. This implies that early-stage patients can reap greater benefits from systemic chemotherapy and presents a potential for conversion therapy. In the process of clinical treatment decision-making, the scope of metastasis should be ascertained prior to commencing systemic chemotherapy. For patients with solely intrahepatic lesions, systemic chemotherapy can be given priority; for those with extrahepatic metastases, local control measures ought to be integrated.

Notably, during the course of whole-body chemotherapy, through the analysis of tumor sizes, it was observed that tumors with a diameter larger than 6 cm exhibited a more favorable treatment response in comparison to smaller tumors. Regarding large-diameter tumors that were responsive to chemotherapy, an earlier shrinkage effect might be more pronounced. However, in the survival analysis of local chemotherapy, no significant treatment disparities were detected, which warrants further investigation. In a subgroup analysis of a phase 3 study, "LEAP-012", of unresectable, nonmetastatic hepatocellular carcinoma, there was a trend toward better survival among patients with a larger tumor burden (number of tumors plus a maximum tumor diameter of more than 6cm) (*18*). For patients with a substantial tumor burden, liver dysfunction may ensue following the progression of TACE, leading to the forfeiture of the opportunity for subsequent treatment (*19,20*). For these patients, it may be more imperative to integrate systemic treatment in advance rather than awaiting the point at which TACE fails to yield benefits before commencing systemic treatment (*21,22*).

In the analysis of adverse events associated with tumor treatment, the primary adverse events in the TACE group were grade 1-2 liver function impairment, and there were no events above grade 3. The main adverse events in the systemic chemotherapy group were bone marrow suppression and weight loss. These conclusions can directly guide clinical pre-treatment. Before TACE, it is necessary to routinely safeguard liver function; before systemic chemotherapy, it is essential to prevent bone marrow suppression and enhance nutritional support, enabling patients to achieve the optimal therapeutic effect and a longer survival time (*23-25*).

Although there have been significant technological advancements and multimodal treatment approaches in clinical practice, the prognosis of ICC remains poor. Local combined with systemic treatment for patients with advanced ICC has gradually shown its effectiveness. Additionally, adjuvant and neoadjuvant chemotherapy, as well as multimodal treatment strategies based on molecular profiling for targeted therapy and immunotherapy, can be applied (*26,27*). These approaches can be discussed in surgical centers and multidisciplinary tumor committees (MDTs) to formulate the best treatment plan for ICC patients (*28*). Moreover, molecular analysis should be conducted for ICC, as approximately 25% of cases have genetic alterations that can be targeted for treatment, providing better evidence support for subsequent research and laying the foundation for subsequent "precision treatment" studies (*29,30*).

This study is a retrospective study with certain limitations, including selection bias. When conducting subgroup analysis for patients who received local chemotherapy, the sample size was relatively small. Moreover, not all patients underwent surgical staging, and there may be potential differences in staging among different treatment groups. However, to minimize this possibility, during the subsequent continuous treatment and follow-up of the patients, we used imaging monitoring to ensure the changes and authenticity of the patient's lesions. Despite these limitations, this study aims to find the best treatment plan for patients with advanced ICC who are not eligible for surgical resection.

In summary, the treatment methods for ICC are constantly being innovated. In this retrospective study, patients who received local chemotherapy, chemotherapy combined with targeted and immunotherapy, and those with lesions confined to the liver all exhibited better survival rates. Additionally, it was observed that when the tumor diameter was at the 6-cm boundary, systemic chemotherapy yielded better results in patients with tumors larger than 6 cm. The results of this study still need to be confirmed in more prospective studies.

Funding: This work was supported by the National Natural Science Foundation of China (No. 82572988).

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Massarweh NN, El-Serag HB. Epidemiology of Hepatocellular Carcinoma and Intrahepatic Cholangiocarcinoma. Cancer

Control. 2017; 24:1073274817729245.
2. Valle JW, Kelley RK, Nervi B, Oh DY, Zhu AX. Biliary tract cancer. Lancet. 2021; 397:428-444.
3. Amini N, Ejaz A, Spolverato G, Kim Y, Herman JM, Pawlik TM. Temporal trends in liver-directed therapy of patients with intrahepatic cholangiocarcinoma in the United States: a population-based analysis. J Surg Oncol. 2014; 110:163-170.
4. de Jong MC, Nathan H, Sotiropoulos GC, *et al.* Intrahepatic cholangiocarcinoma: an international multi-institutional analysis of prognostic factors and lymph node assessment. J Clin Oncol. 2011; 29:3140-3145.
5. Valle J, Wasan H, Palmer DH, Cunningham D, Anthoney A, Maraveyas A, Madhusudan S, Iveson T, Hughes S, Pereira SP, Roughton M, Bridgewater J; ABC-02 Trial Investigators. Cisplatin plus gemcitabine versus gemcitabine for biliary tract cancer. N Engl J Med. 2010; 362:1273-1281.
6. Park J, Kim MH, Kim KP, Park DH, Moon SH, Song TJ, Eum J, Lee SS, Seo DW, Lee SK. Natural History and Prognostic Factors of Advanced Cholangiocarcinoma without Surgery, Chemotherapy, or Radiotherapy: A Large-Scale Observational Study. Gut Liver. 2009; 3:298-305.
7. Okusaka T, Nakachi K, Fukutomi A, Mizuno N, Ohkawa S, Funakoshi A, Nagino M, Kondo S, Nagaoka S, Funai J, Koshiji M, Nambu Y, Furuse J, Miyazaki M, Nimura Y. Gemcitabine alone or in combination with cisplatin in patients with biliary tract cancer: a comparative multicentre study in Japan. Br J Cancer. 2010; 103:469-474.
8. Morizane C, Okusaka T, Mizusawa J, *et al.* Combination gemcitabine plus S-1 versus gemcitabine plus cisplatin for advanced/recurrent biliary tract cancer: the FUGA-BT (JCOG1113) randomized phase III clinical trial. Ann Oncol. 2019; 30:1950-1958.
9. Ioka T, Kanai M, Kobayashi S, *et al.* Randomized phase III study of gemcitabine, cisplatin plus S-1 versus gemcitabine, cisplatin for advanced biliary tract cancer (KHBO1401- MITSUBA). J Hepatobiliary Pancreat Sci. 2023; 30:102-110.
10. Valle JW, Wasan H, Johnson P, *et al.* Gemcitabine alone or in combination with cisplatin in patients with advanced or metastatic cholangiocarcinomas or other biliary tract tumours: a multicentre randomised phase II study - The UK ABC-01 Study. Br J Cancer. 2009; 101:621-627.
11. Valle JW, Wasan H, Lopes A, *et al.* Cediranib or placebo in combination with cisplatin and gemcitabine chemotherapy for patients with advanced biliary tract cancer (ABC-03): a randomised phase 2 trial. Lancet Oncol. 2015; 16:967-978.
12. Oh DY, He AR, Bouattour M, *et al*. Durvalumab or placebo plus gemcitabine and cisplatin in participants with advanced biliary tract cancer (TOPAZ-1): updated overall survival from a randomised phase 3 study. Lancet Gastroenterol Hepatol. 2024; 9:694-704.
13. Kelley RK, Ueno M, Yoo C, *et al.* Pembrolizumab in combination with gemcitabine and cisplatin compared with gemcitabine and cisplatin alone for patients with advanced biliary tract cancer (KEYNOTE-966): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet. 2023; 401:1853-1865.
14. Banales JM, Marin JJG, Lamarca A, *et al*. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. Nat Rev Gastroenterol Hepatol. 2020; 17:557-588.
15. Feng Y, Zhao M, Wang L, Li L, Lei JH, Zhou J, Chen J, Wu Y, Miao K, Deng CX. The heterogeneity of signaling pathways and drug responses in intrahepatic cholangiocarcinoma with distinct genetic mutations. Cell Death Dis. 2024; 15:34.
16. Xiao H, Wang J, Weng Z, *et al*. A histopathology-based artificial intelligence system assisting the screening of genetic alteration in intrahepatic cholangiocarcinoma. Br J Cancer. 2025; 132:195-202.
17. Smolenschi C, Blanc JF, Lancry A, Klajer E, Debaillon-Vesque A, Vantelon JM, Boileve A, Valery M, Hollebecque A, Ducreux M, Decraecker M. Real-world efficacy of zanidatamab in patients with HER2 positive advanced biliary tract cancers. Eur J Cancer. 2025; 222:115432.
18. Kudo M, Ren Z, Guo Y, *et al*. Transarterial chemoembolisation combined with lenvatinib plus pembrolizumab versus dual placebo for unresectable, non-metastatic hepatocellular carcinoma (LEAP-012): a multicentre, randomised, double-blind, phase 3 study. Lancet. 2025; 405:203-215.
19. Jiang N, Zhang Z, Yin X, Qiu H, Yan W, Hao Y, Yang W, Li H, Xu A, Mu K. Systemic chemotherapy plus transarterial chemoembolization versus systemic chemotherapy alone for unresectable intrahepatic cholangiocarcinoma: a multicenter retrospective cohort study. Radiol Med. 2024; 129:631-642.
20. He M, Jiang N, Yin X, Xu A, Mu K. Conventional and drug-eluting beads transarterial chemoembolization in patients with unresectable intrahepatic cholangiocarcinoma: a systematic review and pooled analysis. J Cancer Res Clin Oncol. 2023; 149:531-540.
21. Brown DB, Geschwind JF, Soulen MC, Millward SF, Sacks D. Society of Interventional Radiology position statement on chemoembolization of hepatic malignancies. J Vasc Interv Radiol. 2009; 20(7 Suppl):S317-S323.
22. Gusani NJ, Balaa FK, Steel JL, Geller DA, Marsh JW, Zajko AB, Carr BI, Gamblin TC. Treatment of unresectable cholangiocarcinoma with gemcitabine-based transcatheter arterial chemoembolization (TACE): a single-institution experience. J Gastrointest Surg. 2008; 12:129-137.
23. Luo S, Xiang Z, Li M, Zhao C, Yan H, Huang M. Clinical Effectiveness of Drug-Eluting Microsphere Transcatheter Arterial Chemoembolization Combined with First-Line Chemotherapy as the Initial Treatment for Patients with Unresectable Intrahepatic Cholangiocarcinoma. J Vasc Interv Radiol. 2024; 35:1616-1625.
24. Boehm LM, Jayakrishnan TT, Miura JT, Zacharias AJ, Johnston FM, Turaga KK, Gamblin TC. Comparative effectiveness of hepatic artery based therapies for unresectable intrahepatic cholangiocarcinoma. J Surg Oncol. 2015; 111:213-220.
25. Squires MH, Cloyd JM, Dillhoff M, Schmidt C, Pawlik TM. Challenges of surgical management of intrahepatic cholangiocarcinoma. Expert Rev Gastroenterol Hepatol. 2018; 12:671-681.
26. Wang Y, Shang P, Xu C, Dong W, Zhang X, Xia Y, Sui C, Yang C. Novel genetic alterations in liver cancer distinguish distinct clinical outcomes and combination immunotherapy responses. Front Pharmacol. 2024; 15:1416295.
27. Chen L, Han R, Song T, Song P, Tang W. Perioperative and precision strategies in resectable intrahepatic cholangiocarcinoma. Glob Health Med. 2025; 7:347-351.
28. Heuser C, Diekmann A, Kowalski C, Enders A, Conrad R, Pfaff H, Ansmann L, Ernstmann N. Health literacy

and patient participation in multidisciplinary tumor conferences in breast cancer care: a multilevel modeling approach. BMC Cancer. 2019; 19:330.
29. Yu Y, You Y, Duan Y, *et al*. Multi-omics approaches for identifying the PANoptosis signature and prognostic model *via* a multimachine-learning computational framework for intrahepatic cholangiocarcinoma. Hepatology. 2026; 83:466-483.
30. Zhang XF, Beal EW, Bagante F, *et al.* Early versus late recurrence of intrahepatic cholangiocarcinoma after resection with curative intent. Br J Surg. 2018; 105:848-856.

Received February 13, 2026; Revised April 13, 2026; Accepted April 22, 2026.

[§]These authors contributed equally to this work.
**Address correspondence to:*
Hua Zhang, Department of Hepatobiliary Surgery, The First Affiliated Hospital, Chongqing Medical University, No. 24 Shiyou Rd, Chongqing 400042, China.
E-mail: zhanghuadoctor@163.com

Released online in J-STAGE as advance publication April 26, 2026.

Original Article

DOI: 10.5582/ddt.2026.01005

Hypoalbuminemia and reduced sputum microbiome diversity associated with antibiotic treatment failure in nursing and healthcare-associated pneumonia

Naoki Hosogaya[1,2,§,*], Shoichi Fukui[2,§], Takahiro Takazono[1,3], Koki Fukushima[1], Ryosuke Morio[1], Satoshi Irifune[1], Shimpei Morimoto[2], Nana Nakada[4], Masataka Yoshida[1], Kazuaki Takeda[1], Shotaro Ide[1], Naoki Iwanaga[1], Kazuki Nemoto[5], Koichi Izumikawa[3], Kazuhiro Yatera[5], Katsunori Yanagihara[6], Hiroshi Mukae[1]

[1] Department of Respiratory Medicine, Nagasaki University Hospital, Nagasaki, Japan;
[2] Clinical Research Center, Nagasaki University Hospital, Nagasaki, Japan;
[3] Department of Infectious Diseases, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan;
[4] Health Center, Nagasaki University, Nagasaki, Japan;
[5] Department of Respiratory Medicine, University of Occupational and Environmental Health, Japan, Fukuoka, Japan;
[6] Department of Laboratory Medicine, Nagasaki University Hospital, Nagasaki, Japan.

SUMMARY: Nursing and healthcare-associated pneumonia (NHCAP) pose significant challenges in older populations, yet factors predicting antibiotic treatment failure remain elusive. This exploratory secondary analysis of a multicenter phase IV trial aimed to identify the clinical and microbiome predictors of treatment failure in patients with NHCAP treated with lascufloxacin. Among the 56 evaluable patients (median age 86 years; cured n = 44, not cured n = 12), paired sputum and tongue samples were analyzed using 16S ribosomal RNA gene clone library sequencing. Alpha diversity was assessed using the Shannon index, Simpson index, observed richness, and Pielou's evenness, whereas beta diversity was calculated using Bray-Curtis dissimilarity and visualized by principal coordinate analysis. Serum albumin was significantly lower in not cured patients (3.0 *vs*. 3.5 g/dL, p = 0.0497) and emerged as the strongest predictor of treatment failure in univariate logistic regression (odds ratio 0.18, 95% confidence interval 0.05–0.73, p = 0.016). Sputum Pielou's evenness showed a comparable predictive ability (odds ratio 0.010, p = 0.047). The overall microbiome community composition did not differ according to the outcome. Notably, patients with hypoalbuminemia (< 2.85 g/dL) exhibited significantly reduced sputum alpha diversity (Shannon p = 0.034, Simpson p = 0.025, Pielou's evenness p = 0.010). A simple risk stratification combining hypoalbuminemia and denture use identified a high-risk subgroup with markedly elevated treatment failure rates (75.0% *vs.* 12.5%, p = 0.001). These findings suggest an interconnected pathophysiology linking nutritional status and respiratory microbiome stability in patients with NHCAP. Nutritional status and oral health may be modifiable targets for improving treatment outcomes in high-risk patients.

Keywords: oral microbiome, aspiration pneumonia, hypoalbuminemia, alpha diversity, denture, aged

1. Introduction

Nursing and healthcare-associated pneumonia (NHCAP) represent a significant clinical challenge in Japan's super-aged society (*1*). NHCAP frequently affects older patients with multiple comorbidities and impaired activities of daily living (*2*), with aspiration reported in approximately 68.6% of cases (*3*). Although the 30-day mortality rate is approximately 11.9%, host factors including non-ambulatory status, disturbance of consciousness, and hypoalbuminemia significantly worsen prognosis (*3*).

The oral-lung microbiome axis has emerged as a key factor in the pathogenesis of pneumonia. The oral cavity serves as the primary source of the lung microbiome through continuous microaspiration, even in healthy individuals (*4*). Furthermore, the enrichment of the lung microbiome with oral taxa is associated with Th17-type pulmonary inflammation, suggesting that aspirated oral bacteria actively modulate lung immune responses (*5*). Under pathological conditions, Noguchi *et al*. demonstrated through clone library analysis of bronchoalveolar lavage fluid (BALF) that oral anaerobes were frequently detected in patients with healthcare-associated pneumonia (*6*). Similarly, 16S rRNA gene

analysis of BALF from Japanese patients with hospital-acquired pneumonia revealed that oral streptococci were the most prevalent organisms (23.1%), followed by *Corynebacterium* spp. (11.6%), *Haemophilus* spp. (6.8%), *S. aureus* (6.8%), and *P. aeruginosa* (5.4%) (*7*).

Accumulating evidence supports the clinical significance of oral microbiome in regard to predicting pneumonia outcomes. A prospective cohort study of nursing home residents demonstrated that tongue microbiota dominated by *Prevotella* and *Veillonella* were significantly associated with increased pneumonia mortality, with an adjusted hazard ratio of 13.88 for pneumonia-related death (*8*). Consistent with this finding, a randomized controlled trial in Japanese-assisted living facilities showed that professional oral care reduced pneumonia incidence, fever duration, and mortality (*9*). However, while these studies have focused on the oral environment, pneumonia development, and mortality, the relationship between the oral microbiome composition and response to antibiotic treatment remains elusive.

A recent multicenter phase IV trial reported an overall clinical efficacy rate of 78.6% for lascufloxacin in patients with NHCAP at test-of-cure (TOC) (*10*). The present study was a secondary analysis of a trial aimed at identifying microbiome signatures and clinical factors associated with treatment failure. We hypothesized that specific oral bacterial profiles, particularly those associated with poor oral hygiene and denture use, would be enriched in patients in whom antibiotic treatment fails. Our objectives were to: (1) characterize the clinical predictors of treatment failure; (2) identify the bacterial profiles associated with treatment outcomes; and (3) explore the relationship between the oral microbiome and treatment response.

2. Methods

2.1. Study design and data source

This was a secondary analysis of a multicenter, open-label, phase IV clinical trial evaluating lascufloxacin for NHCAP (*10*). The original trial was conducted at 24 sites in Japan between December 2020 and September 2023. The primary trial enrolled patients with NHCAP, defined according to the Japanese Respiratory Society guidelines (*1*) who were treated with lascufloxacin 75 mg once daily for 7-10 days. This study conformed to the principles of the Declaration of Helsinki (as revised in 2013). The trial was approved by the Clinical Research Review Board in Nagasaki University (approval number: CRB20-023), and written informed consent was obtained from all participants or their legal representatives. This trial is registered with the Japan Registry of Clinical Trials (jRCTs071200066).

2.2. Participants

Of the 75 patients who provided written informed consent to participate in the original trial, 56 met the eligibility criteria and had evaluable outcomes at the TOC visit, comprising the full analysis set (FAS) for the secondary analysis. The primary outcome was the treatment response at TOC, classified as 'cured' (complete resolution of signs and symptoms) or 'not cured' (persistent or worsening of clinical status).

2.3. Clinical assessment

The baseline clinical data included demographics, laboratory parameters, oral health status (remaining teeth, denture use), and aspiration risk. Severity was assessed using the A-DROP (Age, Dehydration, Respiratory failure, Orientation disturbance, low blood Pressure) score (*11*).

2.4. Microbiome sample collection and processing

Paired sputum and tongue dorsum samples were collected at baseline (pretreatment). The sputum samples were collected *via* spontaneous expectoration. Tongue samples were collected by swabbing the dorsal surface of the tongue. The samples were then immediately stored at -80°C until DNA extraction.

2.5. 16S rRNA gene sequencing and clone library analysis

The bacterial flora were analyzed according to the clone library method using amplified fragments of the 16S ribosomal RNA gene, as previously described by Noguchi *et al*. (*6*). Briefly, DNA was extracted from the specimens by vigorous shaking with sodium dodecyl sulfate (final concentration: 3.0%) and glass beads. The 16S rRNA gene was amplified by polymerase chain reaction (PCR) using universal primers. The PCR products were cloned using a TOPO TA cloning kit (*Invitrogen*), and colonies were randomly selected from each clone library for sequencing analysis. Sequences were compared with an in-house database containing the 16S rRNA gene sequences of the type strains using a basic local alignment search tool algorithm. This method demonstrated superior detection of oral bacteria, including streptococci and anaerobes, compared to conventional cultivation methods for healthcare-associated pneumonia (*6*).

2.6. Statistical analysis

Continuous variables were compared using the Wilcoxon rank-sum test and categorical variables were compared using Fisher's exact test. Alpha diversity was assessed using four indices: the Shannon index (*12*), Simpson index (*13*), observed species richness, and Pielou's evenness index (*14*). Beta diversity was calculated using

Bray-Curtis dissimilarity (*15*), visualized by principal coordinate analysis (PCoA) (*16*), and statistically evaluated using permutational multivariate analysis of variance (PERMANOVA) with 999 permutations (*17*).

Given the exploratory nature and limited sample size of this secondary analysis, with only 12 events (not cured) for the minority outcome, the differential abundance of the individual bacterial species was assessed using the Fisher's exact test for detection rates and the Wilcoxon rank-sum test for relative abundance, with *p*-values reported without adjustment for multiple comparisons.

The clinical predictors of treatment failure were evaluated using univariate and multivariable logistic regression analyses. The odds ratios (OR) with 95% confidence intervals (CI) were calculated. The optimal albumin cut-off value was determined using the Youden index (*18*). Statistical significance was defined as a two-sided *p*-value of < 0.05. All analyses were performed using the R version 4.5.0 (R Foundation for Statistical Computing, Vienna, Austria).

3. Results

3.1. Patient characteristics

The FAS comprised 56 patients (cured: *n* = 44; not cured: *n* = 12). The baseline characteristics are presented in Table 1. The median age was 86 years (IQR: 79.5–89), and 58.9% were male. Denture use was common (67.9%), and the median number of remaining teeth was 10.5 (IQR: 0–21.2). Most patients had moderate-severity pneumonia according to the A-DROP score. Two groups showed comparable age, body weight, C-reactive protein, and pneumonia severity, but serum albumin (cured: 3.5 g/dL [3.1–3.7] *vs*. not cured: 3.0 g/dL [2.4–3.5], $p = 0.0497$) was significantly different. Denture use tended to be higher in the not cured patients than in the cured patients

Table 1. Baseline characteristics of the study population

Characteristic	Overall (*N* = 56)	Cured (*n* = 44)	Not Cured (*n* = 12)	*p*-value
Demographics				
Sex, Male/Female	33/23	24/20	9/3	0.322
Age (years)	86.0 (79.5–89.0)	86.0 (79.5–89.0)	85.0 (81.0–89.2)	0.936
Body weight (kg)	46.0 (40.5–49.2)	46.0 (40.4–49.8)	45.8 (40.5–47.9)	0.881
Underlying diseases, *n* (%)	41 (73.2)	32 (72.7)	9 (75.0)	1.000
NHCAP criteria, *n* (%)				
Resident in long-term care facility	23 (41.1)	17 (38.6)	6 (50.0)	0.522
Hospitalization within 90 days	19 (33.9)	14 (31.8)	5 (41.7)	0.516
Elderly requiring nursing care	40 (71.4)	31 (70.5)	9 (75.0)	1.000
Continuous endovascular treatment	5 (8.9)	5 (11.4)	0 (0.0)	0.574
Risk factors for resistant bacteria, n (%)				
Any	23 (41.1)	19 (43.2)	4 (33.3)	0.743
Antimicrobial use within 90 days	15 (26.8)	11 (25.0)	4 (33.3)	0.715
Hospitalization within 90 days	22 (39.3)	17 (38.6)	5 (41.7)	1.000
Nursing home/long-term care	10 (17.9)	6 (13.6)	4 (33.3)	0.196
Enteral feeding	0 (0.0)	0 (0.0)	0 (0.0)	–
A-DROP score	2.0 (1.0–2.0)	2.0 (1.0–2.0)	1.5 (1.0–2.0)	0.778
Risk for aspiration (SSA), *n* (%)				1.000
Concerned	13 (23.2)	10 (22.7)	3 (25.0)	
Not concerned	43 (76.8)	34 (77.3)	9 (75.0)	
Oral status				
Remaining teeth	10.5 (0.0–21.2)	14.0 (0.0–22.5)	3.5 (0.8–16.8)	0.402
Denture use, Yes/No	38/18	27/17	11/1	0.079
Vital signs				
Body temperature (°C)	37.0 (36.6–37.8)	37.0 (36.6–37.7)	37.0 (36.7–37.8)	0.873
Systolic blood pressure (mmHg)	123.0 (109.2–143.5)	123.0 (106.8–145.2)	123.0 (113.8–141.5)	0.952
SpO_2 (%)	96.0 (93.8–97.0)	96.0 (93.0–97.2)	96.0 (94.8–96.2)	0.984
Laboratory data				
White blood cell count ($\times 10^3/\mu L$)	9.4 (7.4–13.0)	10.3 (7.0–13.5)	8.1 (7.6–9.2)	0.353
Neutrophils (%)	80.2 (73.2–84.7)	80.2 (72.3–85.7)	80.6 (78.7–82.3)	0.826
Lymphocytes (%)	11.8 (8.6–15.4)	11.4 (8.3–16.1)	12.2 (10.8–13.7)	0.562
Hemoglobin (g/dL)	11.6 (10.5–12.3)	11.6 (10.5–12.4)	10.6 (10.4–11.8)	0.242
Platelets ($\times 10^4/\mu L$)	21.2 (16.8–24.9)	20.2 (16.5–24.6)	23.4 (21.1–26.7)	0.062
CRP (mg/dL)	4.5 (2.3–10.3)	3.9 (2.3–9.8)	7.5 (3.9–11.8)	0.161
Albumin (g/dL)	3.5 (3.0–3.7)	3.5 (3.1–3.7)	3.0 (2.4–3.5)	0.0497
Total protein (g/dL)	6.8 (6.1–7.2)	6.8 (6.1–7.2)	6.6 (6.3–7.2)	0.718

Values are presented as median (IQR), *n* (%), or *n*/*n*. The Fisher's exact test was used to analyze categorical variables. The Wilcoxon rank-sum test was used to analyze continuous variables. Abbreviations: IQR, interquartile range; NHCAP, nursing and healthcare-associated pneumonia; SSA, Standardized Swallowing Assessment; CRP, C-reactive protein.

(91.7% *vs.* 61.4%, $p = 0.079$).

3.2. Alpha diversity in microbiome and treatment outcome

The overall alpha diversity of the sputum and tongue microbiomes did not differ significantly between the two groups. For sputum samples, Pielou's evenness showed a trend toward lower values in not cured patients ($p = 0.086$), whereas other indices showed no significant associations (Shannon $p = 0.348$, Simpson $p = 0.213$, observed richness $p = 0.787$; Figure 1A). For the tongue samples, no indices showed any significant differences between the two groups (Shannon $p = 0.686$, Simpson $p = 0.745$, observed richness $p = 0.575$, Pielou's evenness $p = 0.836$; Figure 1B).

3.3. Correlation between sputum and tongue alpha diversity

The correlation between sputum and tongue diversity indices was found to be moderate to strong (Figure 2). Shannon index showed significant correlation ($\rho = 0.420$, $p = 0.002$), as did Simpson index ($\rho = 0.339$, $p = 0.012$) and observed richness ($\rho = 0.450$, $p < 0.001$). Pielou's evenness showed a weaker, non-significant correlation ($\rho = 0.250$, $p = 0.069$).

3.4. Beta diversity

Hierarchical clustering heat maps of the bacterial species demonstrated substantial inter-individual variation without clear clustering by outcome (Figure 3A). Beta diversity analysis by PCoA and PERMANOVA showed no significant differences in the community composition between the outcome groups for sputum ($R^2 = 0.0127$, $p = 0.863$) or tongue samples ($R^2 = 0.0147$, $p = 0.700$) (Figure 3B). Paired analysis of sputum versus tongue samples revealed significant differences in the community composition between the sample types (PERMANOVA $R^2 = 0.036$, $p = 0.001$; Figure 3C).

3.5. Exploratory analysis of individual bacterial species

In the exploratory differential abundance analysis, three sputum bacterial species showed significantly different detection rates between the outcome groups (Fisher's exact test, $p < 0.05$, uncorrected; Table 2): *Veillonella*

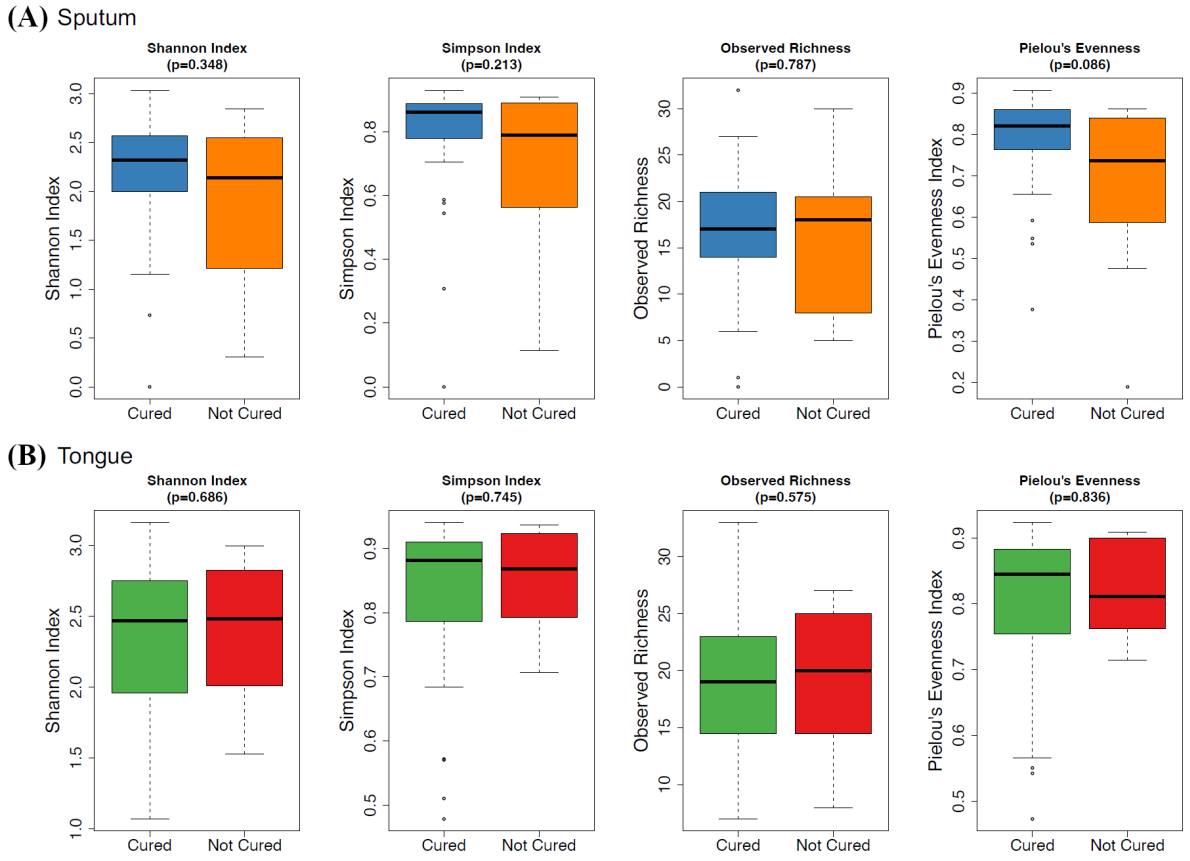

Figure 1. Alpha diversity of sputum and tongue microbiomes by treatment outcome. (**A**) Sputum microbiome alpha diversity indices comparing the cured ($n = 44$) and not cured ($n = 12$) patients. Shannon index ($p = 0.348$), Simpson index ($p = 0.213$), observed species richness ($p = 0.787$), and Pielou's evenness ($p = 0.086$) showed no significant differences, although evenness showed a trend toward lower values in not cured patients. (**B**) Tongue microbiome alpha diversity indices comparing cured and not cured patients. No significant differences were observed for Shannon index ($p = 0.686$), Simpson index ($p = 0.745$), observed richness ($p = 0.575$), or Pielou's evenness ($p = 0.836$). Box plots display median with interquartile range; whiskers extend to 1.5× IQR. P-values were calculated using Wilcoxon rank-sum test.

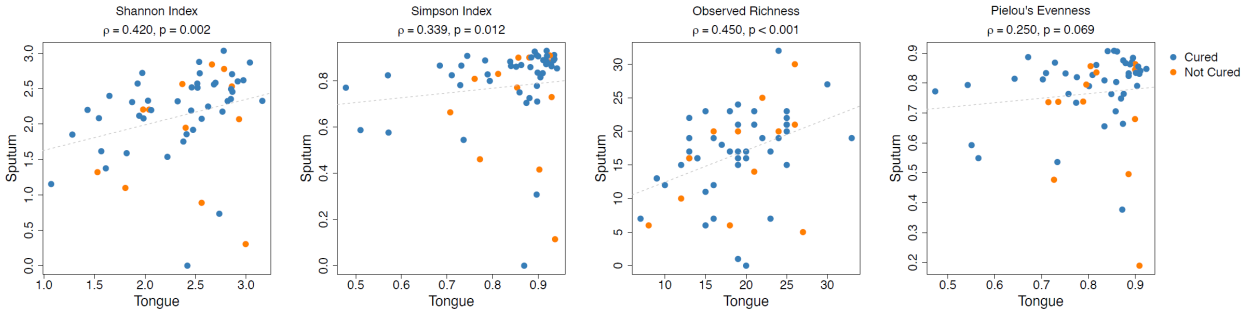

Figure 2. Correlation between sputum and tongue microbiome alpha diversity. Scatter plots showing paired correlations between the sputum (x-axis) and tongue (y-axis) alpha diversity indices for 56 patients. Shannon index (ρ = 0.420, p = 0.002), Simpson index (ρ = 0.339, p = 0.012), and observed species richness (ρ = 0.450, p < 0.001) showed significant positive correlations. Pielou's evenness showed a weaker, non-significant correlation (ρ = 0.250, p = 0.069). Spearman's rank correlation coefficients (ρ) and p-values are displayed for each comparison.

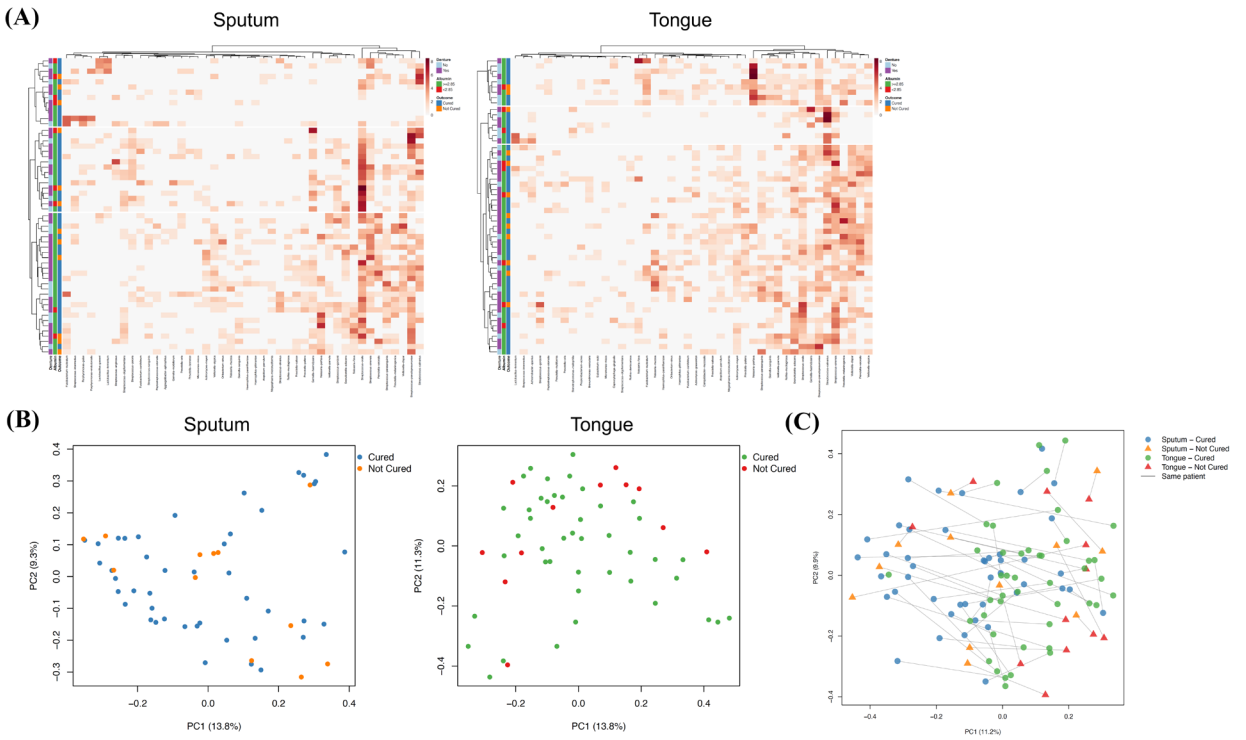

Figure 3. Beta diversity analysis of sputum and tongue microbiome. (**A**) Hierarchical clustering heatmaps of the bacterial species composition in the sputum (left) and tongue (right) samples. Color intensity represents square-root transformed relative abundance (%). Clinical annotations include treatment outcome (Cured/Not Cured), serum albumin status (≥ 2.85 or < 2.85 g/dL), and denture use. Bacterial species detected in fewer than 10% of patients (≤ 5 patients) were excluded from the heatmap visualization. (**B**) Principal coordinates analysis (PCoA) based on Bray-Curtis dissimilarity for the sputum (left) and tongue (right) samples, with points colored by treatment outcome. No significant differences in community composition were observed between outcome groups for sputum (R^2 = 0.0127, p = 0.863) or tongue samples (R^2 = 0.0147, p = 0.700). (**C**) Paired PCoA comparing sputum and tongue samples from the same patients. Lines connect paired samples from the same individual. Community composition was significantly different between sample types (PERMANOVA R^2 = 0.036, p = 0.001).

rodentium (cured 0% *vs.* not cured 16.7%, p = 0.043), *Lautropia mirabilis* (0% *vs.* 16.7%, p = 0.043), and *Campylobacter showae* (0% *vs.* 16.7%, p = 0.043). All three species were detected exclusively in not cured patients and were absent in cured patients. However, these species were detected in only two of the 12 not cured patients (16.7%). *Actinomyces viscosus* showed a similar trend (4.5% *vs.* 25%, p = 0.060). In tongue samples, *Actinomyces graevenitzii* was significantly more frequent in not cured patients (18.2% *vs.* 58.3%, p = 0.010), and *Corynebacterium matruchotii* was exclusively detected in not cured patients (0% *vs.* 16.7%, p = 0.043).

3.6. Clinical predictors of treatment outcome

Receiver operating characteristic (ROC) analysis revealed that serum albumin level had a moderate predictive ability for treatment failure (area under the curve [AUC] = 0.687; Figure 4A). Youden index

Table 2. Significant bacteria associated with treatment outcome (uncorrected $p < 0.05$)

Bacteria	Cured detected, *n* (%)	Not Cured detected, *n* (%)	*p* value for Fisher's exact test	Not Cured median proportion (IQR)	Cured median detected proportion (%, IQR)	Not Cured median detected proportion (%, IQR)	*p* value for Wilcoxon test
Sputum							
Veillonella rodentium	0 (0)	2 (16.7)	0.043	0 (0-0)	NA (NA)	1.26 (NA)	0.007
Lautropia mirabilis	0 (0)	2 (16.7)	0.043	0 (0-0)	NA (NA)	1.35 (NA)	0.007
Campylobacter showae	0 (0)	2 (16.7)	0.043	0 (0-0)	NA (NA)	1.49 (NA)	0.007
Actinomyces viscosus	2 (4.5)	3 (25)	0.060	0 (0-0.35)	1.39 (NA)	1.52 (1.46-2.01)	0.025
Tongue							
Actinomyces graevenitzii	8 (18.2)	7 (58.3)	0.010	1.35 (0-2.41)	2.33 (1.17-2.50)	2.35 (1.61-3.59)	0.004
Corynebacterium matruchotii	0 (0)	2 (16.7)	0.043	0 (0-0)	NA (NA)	5.91 (NA)	0.007

IQR: interquartile range; NA: not applicable (insufficient samples for calculation).

optimization identified 2.85 g/dL as an optimal cutoff (Figure 4B).

3.7. Clinical-microbiome correlations

In the sputum samples, patients with low albumin (< 2.85 g/dL) showed significantly lower alpha diversity compared to those with higher albumin (Figure 5A): Shannon index ($p = 0.034$), Simpson index ($p = 0.025$), and Pielou's evenness ($p = 0.010$). In contrast, denture use was not significantly associated with any sputum alpha diversity measure (Figure 5B). Tongue microbiome diversity did not differ significantly according to the albumin status (Figure 5C) or denture use (Figure 5D).

3.8. Logistic regression for treatment failure

In univariate logistic regression (Table 3), serum albumin was significantly associated with treatment failure (OR = 0.18, 95% CI: 0.05–0.73, $p = 0.016$), as was Pielou's evenness index in sputum (OR = 0.010, 95% CI: < 0.01–0.95, $p = 0.047$). Denture use showed a trend to be associated with treatment failure (OR = 6.93, 95% CI: 0.82–58.58, $p = 0.076$).

Given that our study included only 12 not cured cases (the minority outcome), we included two explanatory variables in the multivariable model (events per variable [EPV] = 6), which is supported by Vittinghoff and McCulloch, who demonstrated that problems are uncommon with 5–9 EPV in their simulation study (*19*). In a two-variable model including both predictors, neither achieved statistical significance (albumin: OR = 0.26, 95% CI: 0.06–1.13, $p = 0.072$; Pielou: OR = 0.050, 95% CI: < 0.01–7.68, $p = 0.246$).

3.9. Risk stratification

The patients were stratified into two risk groups based on a combination of albumin < 2.85 g/dL and denture use (Figure 6). Patients with both risk factors (High Risk group, $n = 8$) showed a markedly elevated treatment failure rate (75.0%, 6/8) when compared to Others (12.5%, 6/48; Fisher's exact test, $p = 0.001$).

4. Discussion

In this exploratory secondary analysis of the lascufloxacin phase IV trial, we identified serum albumin level as the strongest predictor of treatment failure in patients with NHCAP. Sputum microbiome evenness also showed predictive potential, whereas the overall microbiome community composition did not differ between outcome groups. Notably, patients with low albumin levels had significantly reduced sputum diversity. A simple risk stratification, combining low albumin levels and denture use, identified a high-risk subgroup with markedly elevated failure rates.

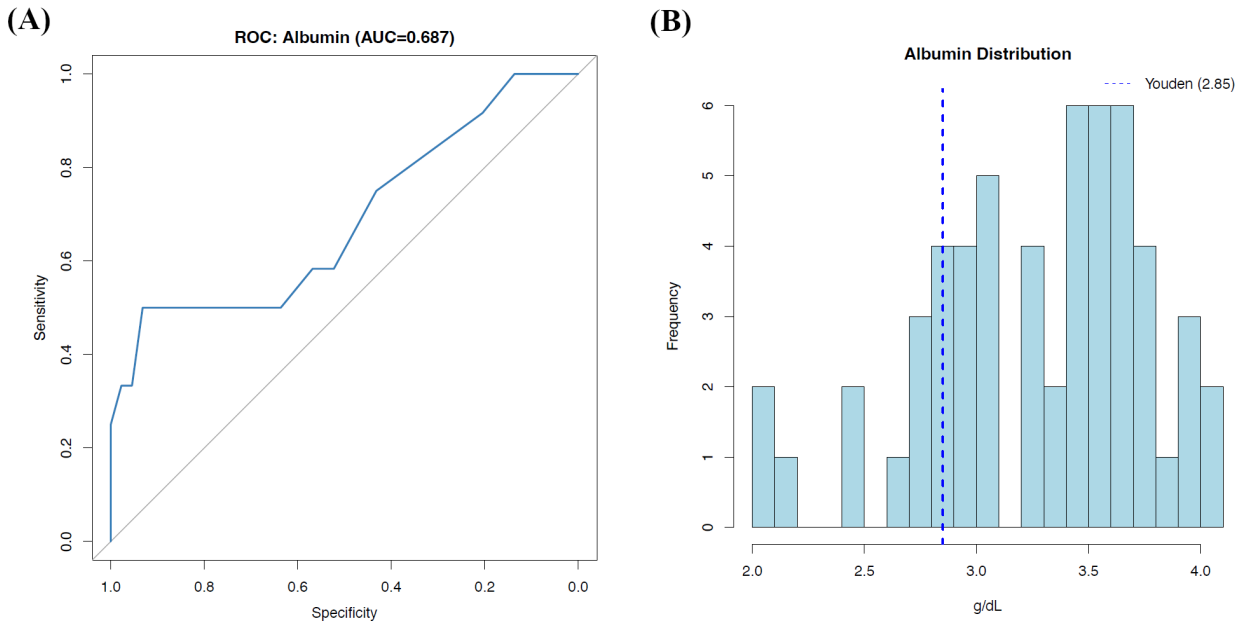

Figure 4. Serum albumin as a predictor of treatment failure. (**A**) Receiver operating characteristic (ROC) curve for serum albumin predicting treatment failure. Area under the curve (AUC) = 0.687. (**B**) Histogram showing the distribution of the baseline serum albumin levels in the study population. The vertical dashed line indicates the optimal cutoff (2.85 g/dL) determined by Youden index.

Serum albumin level emerged as the most robust predictor of treatment failure in univariate analysis. This finding is consistent with a previous large-scale study on community-acquired pneumonia that demonstrated that serum albumin levels at admission were independently associated with ICU admission and 30-day mortality (*20*). Similarly, hypoalbuminemia has been consistently associated with poor prognosis in patients with NHCAP in Japan (*3*).

Sputum Pielou's evenness, a measure of how evenly species are distributed within a community, showed predictive potential comparable to that of albumin, with lower evenness associated with treatment failure. This finding suggests that disruption of the microbiome balance, characterized by the dominance of a few taxa over others, may be more relevant to treatment outcomes than overall species richness. Indeed, different alpha diversity metrics capture distinct aspects of community structure, and evenness indices may reveal dysbiotic states that are not detected by richness alone (*21*).

The relationship between albumin and alpha diversity suggests a shared underlying pathophysiology of NHCAP. Patients with low albumin (< 2.85 g/dL) showed significantly reduced sputum alpha diversity across multiple indices (Shannon $p = 0.034$, Simpson $p = 0.025$, Pielou's evenness $p = 0.010$). Several causal relationships may explain this association: (1) malnutrition may impair immune function and mucosal integrity, promoting oral dysbiosis that extends to the lower respiratory tract; (2) reduced oral microbiome diversity may reflect impaired oral function, affecting nutrient intake; (3) a common upstream factor, such as frailty or systemic inflammation, may affect both; and (4) bidirectional effects may create a vicious cycle of malnutrition, dysbiosis, and impaired host defense. This study precludes the determination of a causal direction, but the observed association suggests that nutritional status may be a key determinant of respiratory microbiome stability in aged patients with pneumonia.

Bacterial species showing exploratory enrichment in patients with treatment failure share the biological characteristics relevant to oral dysbiosis. In sputum, *Veillonella rodentium*, *Lautropia mirabilis*, and *Campylobacter showae* were exclusively detected in not cured patients. *Veillonella* species were highly abundant in both supra- and sub-gingival biofilms (*22*). *Lautropia mirabilis* is a Gram-negative motile coccus originally isolated from the gingival margin and dental plaque of humans; its capacity for polysaccharide production suggests its possible participation in dental plaque formation (*23*). *Campylobacter showae* has also been isolated from periodontitis sites (*24*). However, the detection rate of these bacteria was found to be only 16.7% (2 of 12 not cured patients), severely limiting the reliability of these associations. These findings should therefore be regarded as preliminary exploratory observations, and the potential link between these specific bacterial species and treatment failure requires validation in larger and adequately powered studies.

Denture use showed a trend to be associated with treatment failure (91.7% *vs.* 61.4%, $p = 0.079$). Denture biofilms harbor complex microbial communities that may increase the aspiration risk (*25*), and poor oral hygiene in the aged population is associated with respiratory pathogen colonization (*26*). Interestingly, denture use was not significantly associated with sputum microbiome

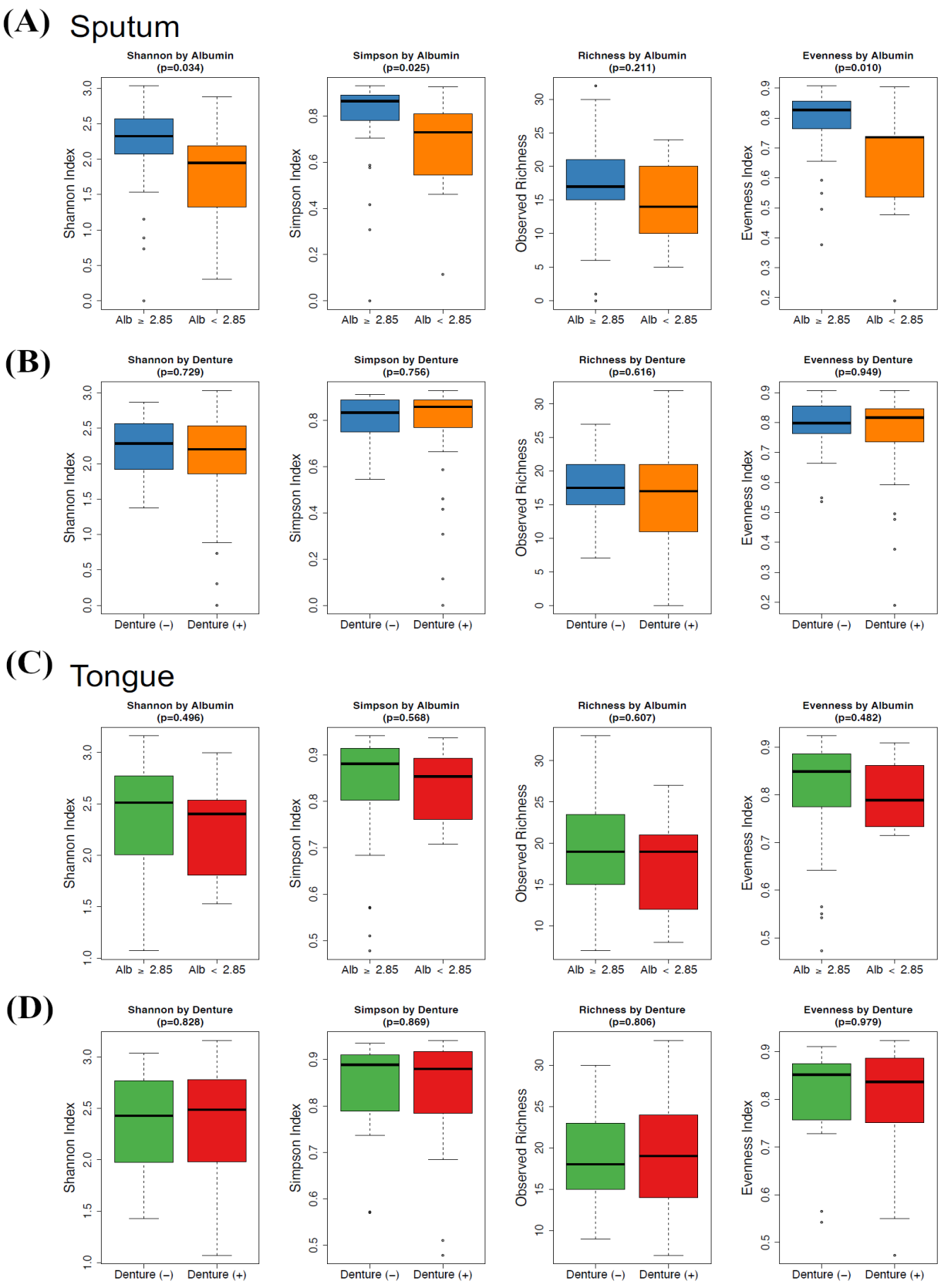

Figure 5. Alpha diversity of sputum and tongue microbiome stratified by albumin status and denture use. (**A**) Sputum microbiome alpha diversity by albumin status. Patients with low albumin (< 2.85 g/dL) showed significantly lower Shannon index (p = 0.034), Simpson index (p = 0.025), and Pielou's evenness (p = 0.010) compared to those with albumin ≥ 2.85 g/dL. (**B**) Sputum microbiome alpha diversity by denture use. No significant associations were observed. (**C**) Tongue microbiome alpha diversity by albumin status. No significant differences were detected. (D) Tongue microbiome alpha diversity by denture use. No significant associations were observed. P-values were calculated using Wilcoxon rank-sum test.

Table 3. Logistic regression analysis for treatment failure

Variable	Unadjusted			Adjusted		
	OR	95% CI	*P*	OR	95% CI	*P*
Serum albumin, g/dL	0.18	0.05–0.73	0.016	0.26	0.06–1.13	0.072
Pielou's evenness (sputum)	0.010	< 0.01–0.95	0.047	0.050	< 0.01–7.68	0.246
Denture use	6.93	0.82–58.58	0.076			
Simpson index (sputum)	0.11	0.01–2.10	0.141			
Shannon index (sputum)	0.54	0.21–1.37	0.192			
Sex, male	2.50	0.60–10.50	0.211			
CRP, mg/dL	1.05	0.95–1.16	0.346			

Abbreviations: OR, odds ratio; CI, confidence interval; CRP, C-reactive protein. Adjusted model includes serum albumin and Pielou's evenness.

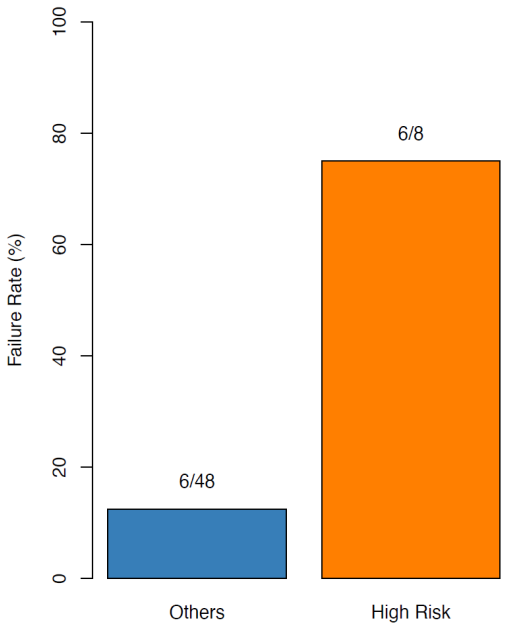

Figure 6. Treatment failure rates by risk stratification. Patients with both risk factors (albumin < 2.85 g/dL and denture use; High Risk group, $n = 8$) showed significantly higher treatment failure rate (75.0%, 6/8) when compared to Others (12.5%, 6/48; Fisher's exact test $p = 0.001$).

diversity, suggesting that the effect of dentures may operate through mechanisms other than community composition, such as increased bacterial load or impaired mucociliary clearance. The combination of denture use and low albumin levels, reflecting both increased bacterial exposure and impaired host defense, created a high-risk phenotype with a significantly elevated treatment failure rate (75.0% *vs*. 12.5%, $p = 0.001$), providing a simple and clinically applicable risk stratification tool.

This study has several limitations that warrant consideration. First, this was a secondary analysis of a single-arm trial without a control group, which limited the causal inference. Second, the sample size was modest ($n = 56$, with only 12 events for the minority outcome), limiting the statistical power for multivariable analysis; the multivariable logistic regression model with two variables was highly unstable, and when both albumin and Pielou's evenness were included in the model, neither reached significance ($p = 0.072$ and $p = 0.246$, respectively). This loss of significance in the multivariable model, despite significance in the univariate analyses, reflects insufficient statistical power to determine whether the two factors act independently or are interrelated. The findings from this study should be regarded as hypothesis-generating rather than confirmatory. Third, the sputum may not accurately represent the microbiome of the lower respiratory tract. Fourth, the 16S rRNA gene clone library method used in this study has an inherently low throughput, with approximately 100 clones sequenced per sample. This technique can detect the most dominant bacterial species but may overlook low-abundance bacteria that could have important biological functions. Consequently, the microbiome profiles described here represent only the most prevalent taxa and the true diversity of the respiratory and oral microbiomes may be underestimated. Fifth, the cross-sectional design precluded the determination of the causal direction between albumin and microbiome diversity.

In conclusion, the serum albumin level was the strongest predictor of NHCAP treatment failure, with sputum microbiome evenness showing comparable predictive ability. This suggests an interconnected pathophysiology linking the nutritional status and respiratory microbiome stability. A simple risk stratification combining low albumin (< 2.85 g/dL) and denture use identified patients at markedly elevated risk. These findings suggest that nutritional optimization and oral health interventions may improve the outcomes of patients with high-risk of NHCAP. Larger longitudinal prospective studies are needed to validate these findings and clarify causal relationships.

Acknowledgements

We thank the patients and families who participated in the original clinical trial, the investigators at all participating sites for providing access to the trial data, and Kyorin Pharmaceutical Co., Ltd. for the funding support.

Funding: The original clinical trial was supported by

Kyorin Pharmaceutical Co., Ltd. This secondary analysis received no additional funding.

Conflict of Interest: Hiroshi Mukae received lecture honoraria and research grants from KYORIN Pharmaceutical Co., Ltd. outside of the submitted work.

References

1. Kohno S, Imamura Y, Shindo Y, Seki M, Ishida T, Teramoto S, Kadota J, Tomono K, Watanabe A. Clinical Practice Guidelines for Nursing- and Healthcare-associated Pneumonia (NHCAP) [Complete translation]. Respir Investig. 2013; 51:103-126.
2. Mukae H, Iwanaga N, Horita N, *et al*. The JRS guideline for the management of pneumonia in adults 2024. Respir Investig. 2025; 63:811-828.
3. Imamura Y, Miyazaki T, Watanabe A, Tsukada H, Nagai H, Hasegawa Y, Tomono K, Ito I, Teramoto S, Ishida T, Kadota J, Kohno S, Mukae H. Prospective multicenter survey for Nursing and Healthcare-associated Pneumonia in Japan. J Infect Chemother. 2022; 28:1125-1130.
4. Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, Beck JM, Curtis JL, Huffnagle GB. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. mBio. 2015; 6:1-10.
5. Segal LN, Clemente JC, Tsay JCJ, *et al*. Enrichment of the lung microbiome with oral taxa is associated with lung inflammation of a Th17 phenotype. Nat Microbiol. 2016; 1:16031.
6. Noguchi S, Mukae H, Kawanami T, Yamasaki K, Fukuda K, Akata K, Ishimoto H, Taniguchi H, Yatera K. Bacteriological Assessment of Healthcare-Associated Pneumonia Using a Clone Library Analysis. PLoS ONE. 2015; 10:e0124697.
7. Nemoto K, Yatera K, Akata K, Ikegami H, Yamasaki K, Hata R, Naito K, Noguchi S, Kawanami T, Fukuda K, Mukae H. Comparative study of bacterial flora in bronchoalveolar lavage fluid of pneumonia patients based on their pneumonia subtypes and comorbidities using 16S ribosomal RNA gene analysis. J Infect Chemother. 2022; 28:1402-1409.
8. Kageyama S, Takeshita T, Furuta M. Relationships of variations in the tongue microbiota and pneumonia mortality in nursing home residents. J Gerontol A Biol Sci Med Sci. 2018; 73:1097-1102.
9. Yoneyama T, Yoshida M, Ohrui T, *et al*. Oral care reduces pneumonia in older patients in nursing homes. J Am Geriatr Soc. 2002; 50:430-433.
10. Takazono T, Hosogaya N, Fukushima K, *et al*. Efficacy and safety of lascufloxacin for nursing- and healthcare-associated pneumonia: A single-arm, open-label clinical trial. J Infect Chemother. 2024; 30:597-602.
11. Miyashita N, Matsushima T, Oka M; Japanese Respiratory Society. The JRS guidelines for the management of community-acquired pneumonia in adults: an update and new recommendations. Intern Med. 2006; 45:419-428.
12. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948; 27:379-423.
13. Simpson EH. Measurement of diversity. Nature. 1949; 163:688.
14. Pielou EC. The measurement of diversity in different types of biological collections. J Theor Biol. 1966; 13:131-144.
15. Bray JR, Curtis JT. An ordination of the upland forest communities of southern Wisconsin. Ecological monographs. 1957; 27:326-349.
16. Gower JC. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika. 1966; 53:325-338.
17. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 2001; 26:32-46.
18. Youden WJ. Index for rating diagnostic tests. Cancer. 1950; 3:32-35.
19. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and cox regression. Am J Epidemiol. 2007; 165:710-718.
20. Viasus D, Garcia-Vidal C, Simonetti A, Manresa F, Dorca J, Gudiol F, Carratalà J. Prognostic value of serum albumin levels in hospitalized adults with community-acquired pneumonia. J Infect. 2013; 66:415-423.
21. Cassol I, Ibañez M, Bustamante JP. Key features and guidelines for the application of microbial alpha diversity metrics. Sci Rep. 2025; 15:622.
22. Zhou P, Manoil D, Belibasakis GN, Kotsakis GA. Veillonellae: Beyond Bridging Species in Oral Biofilm Ecology. Front Oral Health. 2021; 2:774115.
23. Gerner-Smidt P, Keiser-Nielsen H, Dorsch M, Stackebrandt E, Ursing J, Blom J, Christensen AC, Christensen JJ, Frederiksen W, Hoffmann S, Holten-Andersen W, Ying YT. *Lautropia mirabilis* gen. nov., sp. nov., a Gram-negative motile coccus with unusual morphology isolated from the human mouth. Microbiology. 1994; 140:1787-1797.
24. Macuch PJ, Tanner AC. Campylobacter species in health, gingivitis, and periodontitis. J Dent Res. 2000; 79:785-792.
25. O'Donnell L, Smith K, Williams C. Dentures are a reservoir for respiratory pathogens. J Prosthodont. 2016; 25:99-104.
26. Khadka S, Khan S, King A, Goldberg LR, Crocombe L, Bettiol S. Poor oral hygiene, oral microorganisms and aspiration pneumonia risk in older people in residential aged care: a systematic review. Age Ageing. 2021; 50:81-87.

Received January 26, 2026; Revised March 14, 2026; Accepted March 27, 2026.

[§]These authors contributed equally to this work.
**Address correspondence to:*
Naoki Hosogaya, Clinical Research Center, Nagasaki University Hospital, 1-7-1 Sakamoto, Nagasaki 852-8501, Japan.
E-mail: nhosogaya@nagasaki-u.ac.jp

Released online in J-STAGE as advance publication March 31, 2026.

Original Article

DOI: 10.5582/ddt.2026.01008

Efficacy and safety of filgotinib versus tocilizumab in active rheumatoid arthritis: A randomized, open-label, multicenter study with clinical and musculoskeletal ultrasound evaluation (TRANSFORM study)

Toshimasa Shimizu[1,2], Shin-ya Kawashiri[1,3,*], Tomohiro Koga[1], Rieko Kiya[2], Michiko Morita[2], Rina Kawasaki[2], Shohei Kuroda[2], Shigeki Tashiro[2], Shuntaro Sato[2], Moemi Yabe[4], Kenta Misaki[4], Shunichiro Hanai[5], Daiki Nakagomi[5], Michihiro Ogasawara[6], Naoto Tamura[6], Rina Watanabe[7], Hiroshi Kanazawa[7], Tatsuya Atsumi[8], Yukitaka Ueki[9], Tadashi Okano[10], Takahisa Suzuki[11], Hirokazu Takaoka[12], Hiroaki Hamada[13], Toshihiko Hidaka[14], Shunsuke Furuta[15], Naoki Hosogaya[2], Hiroshi Yamamoto[2], Atsushi Kawakami[1]

[1] Department of Immunology and Rheumatology, Division of Advanced Preventive Medical Sciences, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan;
[2] Clinical Research Center, Nagasaki University Hospital, Nagasaki, Japan;
[3] Department of Community Medicine, Division of Advanced Preventive Medical Sciences, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan;
[4] Department of Rheumatology, Kita-Harima Medical Center, Hyogo, Japan;
[5] Department of Rheumatology, University of Yamanashi Hospital, Yamanashi, Japan;
[6] Department of Internal Medicine and Rheumatology, Juntendo University School of Medicine, Tokyo, Japan;
[7] Department of Rheumatology, Aomori Prefectural Central Hospital, Aomori, Japan;
[8] Department of Rheumatology, Endocrinology and Nephrology, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, Sapporo, Japan;
[9] Department of Rheumatology, Hakujyujikai Sasebo Chuo Hospital, Sasebo, Japan;
[10] Center for Senile Degenerative Disorders (CSDD), Osaka Metropolitan University Graduate School of Medicine, Osaka, Japan;
[11] Department of Rheumatology, Japanese Red Cross Nagasaki Genbaku Hospital, Nagasaki, Japan;
[12] Section of Internal Medicine and Rheumatology, Kumamoto Shinto General Hospital, Kumamoto, Japan;
[13] Department of Orthopedic Surgery, NHO Miyakonojo Hospital, Miyazaki, Japan;
[14] Institute of Rheumatology, Miyazaki-Zenjinkai Hospital, Miyazaki, Japan;
[15] Department of Allergy and Clinical Immunology, Chiba University Hospital, Chiba, Japan.

SUMMARY: Janus kinase (JAK) and interleukin-6 (IL-6) inhibitors are therapeutic options for patients with rheumatoid arthritis (RA) with inadequate response to conventional synthetic disease-modifying antirheumatic drugs (csDMARDs); however, no randomized controlled trial has compared their efficacy and safety. Since both act through the JAK–signal transducer and activator of transcription pathway, a comparative evaluation is warranted. We conducted a prospective, randomized, open-label trial at 55 centers in Japan, randomizing patients with active RA despite csDMARD therapy in a 1:1 ratio to receive 200 mg/day filgotinib or subcutaneous tocilizumab as monotherapy; the primary endpoint was American College of Rheumatology (ACR) 50 at week 12, and secondary endpoints included clinical disease activity indices, musculoskeletal ultrasonography scores, patient-reported outcomes, and serum biomarkers through 52 weeks. Twenty-six patients were enrolled (13 per group) before study termination due to insufficient recruitment, and descriptive analyses were performed. At week 12, ACR50 was achieved in 38.5% (5/13) patients in the filgotinib group and 46.2% (6/13) in the tocilizumab group (risk difference: −7.69%; 95% confidence interval: −42.26 to 28.8). Both groups showed early and sustained improvements in disease activity from week 2. The improvement in patient global assessment scores was greater with filgotinib at week 2 but the difference diminished thereafter, and serum IL-6 level decreased with filgotinib but increased with tocilizumab. Four serious adverse events occurred with filgotinib, including infections and cardiac events. Because this study was underpowered and the analysis was descriptive, larger studies are needed to confirm these findings and define optimal use. (The study was registered in the Japan Registry of Clinical Trials (*https://jrct.niph.go.jp*) as jRCTs071200107 and in ClinicalTrials.gov as NCT05090410.)

Keywords: rheumatoid arthritis, filgotinib, JAK inhibitor, tocilizumab, musculoskeletal ultrasound, biomarker

1. Introduction

Rheumatoid arthritis (RA) is a chronic systemic inflammatory disease affecting the synovial joints (*1*). Uncontrolled disease activity in RA may lead to joint destruction and deformities, impairing the quality of life. Therefore, controlling the disease activity using a treat-to-target strategy is required to prevent joint destruction (*2*).

The standard initial treatment in patients with active RA is conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), including methotrexate (MTX); however, a considerable proportion of patients are refractory to treatment with csDMARDs. The choice of DMARDs in the second phase of treatment is crucial for patients with inadequate or intolerant csDMARDs (*3,4*). Biological DMARDs (bDMARDs), primarily used in the second phase after an inadequate response to csDMARDs, provide better clinical outcomes, including clinical remission in patients with RA. Recently, janus kinase (JAK) inhibitors have emerged as the second choice of treatment for patients with RA with an inadequate response or intolerance to csDMARDs (*3,4*).

Proinflammatory cytokines — overproduced and overexpressed — such as interleukin-6 (IL-6), bind to their receptors to activate the JAK-signal transducer and activator of transcription (STAT) signaling pathways, which are implicated in the pathogenesis of RA (*5*). JAK inhibitors effectively suppress disease activity by inhibiting the JAK-STAT signaling pathways.

Filgotinib is a preferential JAK1 inhibitor developed by Gilead (Foster City, CA, USA). In previous studies, approximately 50% of patients with RA receiving filgotinib achieved clinical remission after an inadequate response to csDMARDs (*6,7*). In addition, the effects of JAK inhibitors, including those of filgotinib, are non-inferior or superior to those of tumor necrosis factor (TNF) inhibitors in patients with active RA and an inadequate response to MTX (*7-10*); however, currently, no head-to-head comparison between JAK and IL-6 inhibitors has been performed. IL-6 inhibitors indirectly suppress the JAK-STAT pathway by inhibiting IL-6 signaling. IL-6 plays a key role in activating JAK1-dependent pathways (*5*). Among the currently available JAK inhibitors, filgotinib exhibits a relatively high selectivity for JAK1. Therefore, a comparative evaluation of the efficacy and safety profiles of filgotinib and IL-6 inhibitors is warranted.

Clinical remission can be achieved in a substantial proportion of patients with RA through the introduction of JAK inhibitors or bDMARDs. However, even in patients who achieve clinical remission, musculoskeletal ultrasonography (MSUS) may reveal residual synovitis, an important finding that predicts subsequent joint damage and clinical relapse (*11-14*). MSUS is widely used for evaluating disease activity in RA as a noninvasive, objective, relatively inexpensive, and repeatable imaging modality (*15,16*). Incorporating MSUS alongside clinical disease activity indices provides a more comprehensive assessment of treatment response at the joint level.

This study aimed to evaluate the non-inferiority of filgotinib monotherapy over IL-6 inhibitor monotherapy in patients with RA with an inadequate response to csDMARDs. In addition, we evaluated changes in disease activity using both MSUS and clinical disease activity indices to achieve a more precise assessment in this population. We also performed a comprehensive analysis of serum biomarkers, including a wide range of cytokines and chemokines. However, this study was terminated prematurely because patient enrollment did not progress sufficiently to achieve the target sample size. Given that the planned sample size was not reached, the prespecified non-inferiority analysis was not performed; instead, we limited the analyses to descriptive statistics and estimation-focused analyses.

2. Materials and Methods

2.1. Study design

This prospective, randomized, open-label, two-arm, and interventional clinical trial was conducted at 55 centers across Japan (Supplementary Table S1, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*). The study was registered in the Japan Registry of Clinical Trials (*https://jrct.niph.go.jp*) as jRCTs071200107 and in ClinicalTrials.gov as NCT05090410. It was approved by the certified review board of Nagasaki University (reference number, CRB20-026). Written informed consent was provided by patients before enrollment. The study was conducted in accordance with the principles of the Declaration of Helsinki (*17*), Clinical Trials Act (since February 2019), Act on the Protection of Personal Information and related regulatory notifications, and approved study protocol.

2.2. Changes from the previously published protocol

The study protocol has been previously published (*18*). This study was initiated on March 3, 2021. However, due to delays in enrolling the target number of patients, the study protocol was amended in November 2022—at which time 18 patients had enrolled—to facilitate further recruitment. Specifically, 1) the inclusion criteria were changed from patients with an inadequate response to MTX to patients with an inadequate response to csDMARDs; 2) in the IL-6 inhibitor group, initially including only subcutaneous tocilizumab, the options were expanded to include intravenous tocilizumab and subcutaneous sarilumab; 3)

medications not permitted for previous use were revised from all JAK inhibitors and IL-6 inhibitors to filgotinib only; and 4) the week 8 study visit was excluded. These protocol amendments may have influenced some results of the study, compared to the prospective results with the initial plan. However, the study was continued as a randomized trial, maintaining its primary objective of comparing the efficacy of filgotinib and IL-6 inhibitors in patients with active RA. Given this consistency in the main goal of the study, we decided to proceed the trial with the revised protocol. This manuscript presents the final trial report of the study previously described in the published protocol paper, including the results obtained with the original as well as the revised protocols. The final revised methods are described below.

2.3. Patients

The inclusion criteria were as follows: 1) age ≥ 18 years; 2) diagnosis of RA based on the American College of Rheumatology (ACR) /European League Against Rheumatism 2010 RA Classification Criteria (*19*); 3) at least moderate disease activity, defined as a Disease Activity Score 28 (DAS28) – erythrocyte sedimentation rate (ESR) ≥ 3.2 at eligibility evaluation; 4) history of csDMARDs treatment for ≥ 8 weeks before providing consent, including ≥ 4 weeks at the same doses of csDMARDs; and 5) ability and willingness to provide written informed consent and comply with the study protocol requirements.

The exclusion criteria were as follows: 1) concurrent use of a glucocorticoid equivalent to > 5 mg/day of prednisolone; 2) contraindication for filgotinib, tocilizumab or sarilumab; 3) previous use of filgotinib; 4) treatment with a glucocorticoid and csDMARD and change of dose within 4 weeks before providing consent; 5) treatment with a bDMARD or a biosimilar DMARD (such as infliximab, biosimilar of infliximab, adalimumab, biosimilar of adalimumab, golimumab, certolizumab pegol, or abatacept) within 8 weeks before providing consent; 6) treatment with a TNF inhibitor (such as etanercept or biosimilar of etanercept) within 4 weeks before providing consent; 7) use of a prohibited drug or therapy, other than the agents listed, within 4 weeks before providing consent; 8) complications causing musculoskeletal disorders other than RA (such as ankylosing spondyloarthritis, reactive arthritis, psoriatic arthritis, crystal-induced arthritis, systemic lupus erythematosus, systemic sclerosis, inflammatory myopathy, or mixed connective tissue disease); 9) current pregnancy, breastfeeding, or nonadherence to a medically approved contraceptive regimen during and 12 months after the study period; or 10) inappropriateness for study inclusion as determined by the investigator.

2.4. Intervention

Patients were randomized in a 1:1 ratio to receive either 200 mg/day filgotinib or an IL-6 inhibitor (intravenous tocilizumab 8 mg/kg every 4 weeks, subcutaneous tocilizumab 162 mg/biweekly, or subcutaneous sarilumab 200 mg/biweekly) switched from csDMARDs throughout the study period. Randomization was performed using the minimization method, stratified by RA disease duration (< 2 years and ≥ 2 years), disease activity (DAS28-ESR > 5.1 and ≤ 5.1), and body weight (< 60 kg and ≥ 60 kg). Investigators used an electronic data capture (EDC) system (DATATRAK EDC; Fountayn, Datatrak International, Inc., Beachwood, OH, USA), built by an independent data manager, to assign patients to the filgotinib or IL-6 inhibitor group in a 1:1 ratio using computer-generated random numbers automatically.

Patients assigned to the IL-6 inhibitor group received one of the IL-6 inhibitors — intravenous tocilizumab, subcutaneous tocilizumab, or subcutaneous sarilumab — based on the judgment of the study investigator. Patients with a moderate renal dysfunction (estimated glomerular filtration rate 30–60 mL/min/1.73 m^2) were administered 100 mg/day filgotinib. All patients continued receiving the same doses of glucocorticoids that they were receiving before providing consent throughout the study period. During the study period, the following treatments were prohibited: administration of a bDMARD (except tocilizumab and sarilumab) or JAK inhibitor (except for filgotinib); concomitant use of an immunosuppressant (azathioprine, cyclophosphamide, or cyclosporine) or oral glucocorticoids equivalent to ≥ 5 mg/day of prednisolone, in addition to intra-articular glucocorticoid injections in joints; and nonsteroidal anti-inflammatory drug suppositories. During the study period, the dose of any oral nonsteroidal anti-inflammatory drug was modified only within the range of its approved doses in Japan. Although csDMARDs were prohibited, their addition was allowed if clinical disease activity worsened after 12 weeks. For patients receiving subcutaneous tocilizumab, the administration frequency was adjusted from 162 mg every 2 weeks to once weekly. The flowchart of participant selection is shown in Supplementary Figure S1 (*https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*).

2.5. Patient discontinuation criteria

Participants could be withdrawn prematurely for any of the following reasons: 1) in the filgotinib group, interruption of filgotinib administration for > 7 consecutive days; 2) in the IL-6 inhibitor group, discontinuation of IL-6 inhibitor administration for ≥ 2 consecutive scheduled injections; 3) participant request to withdraw from the trial; 4) participant request to change or discontinue the assigned treatment;

5) adverse events that made continued participation inadvisable; 6) pregnancy; or 7) at the discretion of the principal investigator if continued participation was considered detrimental to the well-being of the participant.

Participants who discontinued were assessed at the time of discontinuation whenever possible, provided the participant cooperated.

2.6. Outcome measurements

Study visits were conducted at baseline and 2, 4, 12, 24, 36, and 52 weeks after the administration of filgotinib or the IL-6 inhibitor. These assessments are shown in Supplementary Table S2 (*https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*). Clinical physicians were blinded to the results of the joint assessments that were performed using MSUS.

Clinical disease activity was evaluated by each attending physician (Japan College of Rheumatology-certified rheumatologists) based on the values of the ACR core set, DAS28-ESR, DAS28-C reactive protein (CRP) (*20*), clinical disease activity index (CDAI), and simplified disease activity index (SDAI) levels. For tender (68) and swollen joints (66), improvement in three of the following five assessments defined the ACR response: 1) patient's global assessment, 2) patient's pain assessment, 3) evaluator's global assessment, 4) Health Assessment Questionnaire-Disability Index (HAQ-DI) (*21*), and 5) CRP or ESR. ACR response rates were defined as ACR20, ACR50, and ACR70 responses, based on an improvement of ≥ 20%, ≥ 50%, or ≥70%, respectively. Each patient's global and pain assessments, and the evaluator's global assessment, were performed on a 0–100-mm visual analog scale (VAS). Patient-reported outcomes (PROs) were evaluated using the duration and severity of morning stiffness, EuroQol 5 dimensions 5-level (EQ-5D-5L), and functional assessment of chronic illness-fatigue (FACIT-F).

Participants underwent MSUS at baseline and 4, 12, 24, 36, and 52 weeks, performed by a Japan College of Rheumatology-certified sonographer. A systematic multiplanar grayscale (GS) and power Doppler (PD) examination of each patient's joint was performed using a multifrequency linear transducer (12–24 MHz). A PD was used depending on the most sensitive Doppler modality on the individual machines. Doppler settings were adjusted at each hospital according to the published recommendations (*22*). During the study, no MSUS settings changes or software upgrades occurred. Joint synovitis was assessed using MSUS in the dorsal views of 22 joints: bilateral wrist, 1st–5th metacarpophalangeal (MCP), interphalangeal (IP), and 2nd–5th proximal interphalangeal (PIP) joints. Each joint was scored for GS and PD on a scale of 0 to 3 in a semiquantitative manner. The sum of the GS or PD scores was considered as the total GS or PD score, respectively. We also assessed the Global Outcome Measures in Rheumatology-European League Against Rheumatism Synovitis Score (GLOESS). GLOESS has been combined with synovial hypertrophy, as shown by GS and PD (*23*).

Radiographic imaging of the bilateral hands (posteroanterior view) and feet (anteroposterior view) was conducted. Trained Japan College of Rheumatology-certified rheumatologists (T.K. and T.S.) evaluated joint damage progression based on the van der Heijde-modified total Sharp score (vdH-mTSS) method, as previously described (*24*), including 16 areas in each hand for erosions and 15 areas for joint space narrowing (*25*).

2.7. Biomarker measurements

Serum concentrations of the following biomarkers were measured: rheumatoid factor (RF) was determined using a latex agglutination turbidimetric immunoassay (LZ test "Eiken" RF) (Eiken Chemical Co., Ltd., Tokyo, Japan). Anti-cyclic citrullinated peptide (CCP) antibodies were quantified using a chemiluminescent immunoassay (STACIA MEBLux test CCP) (Medical & Biological Laboratories Co., Ltd., Tokyo, Japan). Matrix metalloproteinase-3 levels were measured using a latex turbidimetric immunoassay (Panaclear metalloproteinase-3 "Latex") (Sekisui Medical Co., Ltd., Tokyo, Japan).

Multiplex cytokine/chemokine bead assays were performed using diluted serum supernatants and the MILLIPLEX MAP Human Cytokine/Chemokine Magnetic Bead Panel (Merck KGaA, Darmstadt, Germany). Bio-Plex Pro Human Cytokine Assays (Bio-Rad, Hercules, CA, USA) were performed using a Bio-Plex MAGPIX™ Multiplex Reader (Bio-Rad) according to the manufacturer's instructions.

The cytokines/chemokines measured by the bead panel include IL-1α, IL-1β, IL-1 receptor antagonist, IL-2, IL-4, IL-5, IL-6, IL-7, IL-8, IL-10, IL-12 (p40), IL-12 (p70), IL-13, IL-15, IL-17A, IL-17F, IL-18, IL-22, IL-27, interferon-gamma (IFN-γ), IFN-α2, C-X-C motif chemokine ligand 1 (CXCL1) (growth-related oncogene), granulocyte-macrophage colony-stimulating factor, granulocyte colony-stimulating factor, C-X3-C motif chemokine ligand 1 (CX3CL1) (fractalkine), flt-3 ligand, fibroblast growth factor-2, eotaxin, epidermal growth factor, vascular endothelial growth factor, platelet-derived growth factor-AA, soluble CD40 ligand, TNF-α, TNF-β, transforming growth factor-α, C-C motif chemokine ligand (CCL)4 (macrophage inflammatory protein-1β), CCL3 (macrophage inflammatory protein-1α), CCL22 (macrophage-derived chemokine), CCL7 (monocyte chemotactic protein-3), CCL2 (monocyte chemotactic protein-1), CXCL10 (IFN-γ-inducible protein-10), vascular cell adhesion

molecule-1, and intercellular adhesion molecule-1. Serum IL-6 and TNF-α levels were measured using specific enzyme-linked immunosorbent assay kits (R&D Systems, Minneapolis, MN, USA).

2.8. Study endpoints

The primary endpoint was the ACR50 response at week 12. The secondary endpoints of this study were as follows: 1) ACR responses: ACR50 at weeks 2, 4, 24, 36, and 52, and ACR20 and ACR70 at weeks 2, 4, 12, 24, 36, and 52; 2) clinical disease activity: change from baseline in CDAI, SDAI, DAS28-ESR, and DAS28-CRP at weeks 2, 4, 12, 24, 36, and 52; 3) MSUS scores: change from baseline in total PD and GS scores, and GLOESS at weeks 4, 12, 24, 36, and 52; 4) PROs: change from baseline in patient global and pain VAS scores, HAQ-DI, EQ-5D-5L, FACIT-F, and morning stiffness (duration and activity) at weeks 2, 4, 12, 24, 36, and 52; 5) radiographic assessment: change from baseline in vdH-mTSS at weeks 24 and 52; and 6) biomarkers: change from baseline in serum biomarker levels at weeks 2, 4, 12, 24, 36, and 52. In addition, safety endpoints were adverse events (AEs) and study discontinuation rates.

2.9. Statistical analysis

The sample size was determined to ensure a statistical power of 0.80 for the primary analysis. Specifically, simulation data were generated through 4,000 random samplings from a binomial distribution B (n, p) (two samples, 2,000 pairs), and the same analytical procedure as in the primary analysis was applied to these data. The minimum value of n for which non-inferiority was demonstrated in at least 1,600 of 2,000 trials (≥ 80%) was selected. Based on previous research, the success probability p was fixed at 0.40 for both groups (*26*). This simulation indicated that 176 patients per group were required for adequate power. Considering an estimated dropout rate of approximately 12% at the time of primary endpoint assessment, the target enrollment was set at 200 patients per group.

Given that the study did not reach the target sample size, the prespecified non-inferiority analysis was deemed inappropriate. Therefore, a confirmatory hypothesis test was not performed, and analyses focused on estimation rather than hypothesis testing. For the primary endpoint, risk differences and risk ratios with 95% confidence intervals (CIs) were calculated. The 95% CI for the risk difference was estimated using the Mee method (*27*), and that for the risk ratio was estimated using a continuity-corrected approach. For continuous outcomes, medians with interquartile ranges were reported. The probability of study continuation in each group was illustrated using Kaplan–Meier curves. The safety analysis set included all patients who received ≥ 1 dose of the study drug. The full analysis set was defined as all patients in the safety analysis set, randomized to either treatment group, and who had available baseline data for the ACR core set. All statistical analyses were performed using R version 4.4.0 (R Project for Statistical Computing, Vienna, Austria).

3. Results

3.1. Patients' characteristics

Between March 3, 2021, and December 4, 2023, the target sample size was not achieved; 26 patients were enrolled and randomized (13 and 13 to filgotinib and IL-6 inhibitor groups, respectively). In the IL-6 inhibitor group, all patients received subcutaneous

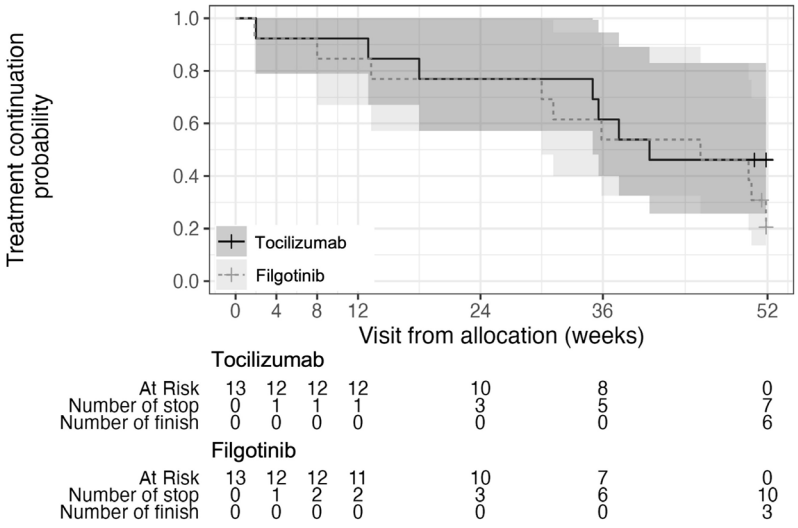

Figure 1. Treatment continuation over 52 weeks. Kaplan–Meier curves show the probability of continuing the assigned study treatment after randomization in the filgotinib (*n* = 13) and tocilizumab (*n* = 13) groups. Shaded areas indicate 95% confidence intervals.

Table 1. Baseline characteristics

Characteristics	Tocilizumab (*n* = 13)	Filgotinib (*n* = 13)
Age (years)[a]	65 (60, 71)	74 (63, 78)
Female[b]	13 (100)	10 (76.9)
Height (cm)[a]	153.6 (152.1, 154.5)	152.6 (150.0, 158.5)
Weight (kg)[a]	51 (46, 63)	48 (42, 65)
Disease duration (years)[a]	4 (0, 13)	2 (0, 5)※
Rheumatoid factor positive history[b]	12 (92.3)	11 (84.6)
Baseline rheumatoid factor (IU/mL)[a]	49 (30, 83)	76 (32, 191)
Anti CCP antibody-positive history[b]	12 (92.3)	8 (61.5)
Baseline anti CCP antibody (U/mL)[a]	92 (15, 143)	12 (6, 373)
Smoking history[b]		
Former smoker	1 (7.7)	2 (15.4)
Current smoker	1 (7.7)	1 (7.7)
Current csDMARDs		
Type of csDMARDs[b]		
MTX	12 (92.3)	13 (100)
Non-MTX	1 (7.7)	0 (0)
MTX dose (mg/week)[a]	9 (7.5, 10.5)※	8 (8, 12)
Pretreatment for rheumatoid arthritis		
Biologics agents[b]	5 (38.5) Infliximab 1, adalimumab 1, etanercept 1, tocilizumab 1, and otilimab 1	3 (23.1) Infliximab 2, and ozoralizumab 1
Concomitant medications		
Glucocorticoid[b]	2 (15.4)	0 (0)
Disease activity[a]		
Tender joint count of 68 joints	8 (7, 13)	8 (3, 22)
Swollen joint count of 66 joints	7 (3, 10)	7 (3, 9)
ESR (mm/h)	40 (12, 44)	38 (17, 66)
CRP (mg/dL)	0.53 (0.12, 0.94)	0.97 (0.5, 1.81)
CDAI	20.8 (18.6, 33)	21.5 (17.7, 28.7)
SDAI	22.2 (18.7, 33)	22.6 (18, 31.3)
DAS28-ESR	5.18 (4.72, 5.94)	5.35 (3.98, 6.33)
DAS28-CRP	4.51 (3.77, 4.93)	4.66 (3.57, 5.57)
Total GS score	13 (8, 15)	12 (7, 24)
Total PD score	5 (1, 10)	7 (4, 19)
GLOESS	14 (9, 15)	13 (7, 29)
Patient global VAS (scale: 0–100)	51.5 (40, 68)	52.9 (23, 80)
Patient pain VAS (scale: 0–100)	53.5 (40, 73)	53.8 (22.1, 80)
Morning stiffness severity (scale: 0–100)	50 (39, 67)	63 (15, 85)
Duration of morning stiffness, minutes	25 (10, 30)	90 (10, 360)
EQ-5D-5L	0.749 (0.636, 0.844)	0.758 (0.653, 0.895)
FACIT-F	39 (34, 47)	38 (33, 41)
HAQ-DI	1 (0.5, 1.25)	0.625 (0.375, 1.62)
vdH-TSS	8.5 (4.5, 31.5)	14 (5, 24.5)

Values are presented as median (IQR) for continuous variables[a] and *n* (%) for categorical variables[b]. CCP, cyclic citrullinated peptide; CDAI,Clinical Disease Activity Index; CRP, C-reactive protein; csDMARDs, conventional synthetic disease-modifying antirheumatic drugs; DAS28, Disease Activity Score-28; ESR, erythrocyte sedimentation rate; EQ-5D-5L, EuroQol 5 Dimension 5-Level; FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; GLOESS, Global the global Outcome Measures in Rheumatology-European League Against Rheumatism Synovitis Score; GS, grayscale; HAQ-DI, Health Assessment Questionnaire-Disability Index; IQR, interquartile range; MTX, methotrexate; PD, power Doppler; RF, rheumatoid factor; SDAI, Simplified Disease Activity Index; VAS, Visual Analog Scale; vdH-TSS, van der Heijde-modified total Sharp score. ※*n* = 12 (1 missing).

tocilizumab 162 mg/biweekly. Therefore, this group would hereafter be referred as the tocilizumab group. All 26 patients were included in both the full and safety analyses sets. During the follow-up period, 10 patients in the filgotinib group and 7 in the tocilizumab group discontinued the study (Figure 1 and Supplementary Figure S1, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*). In the filgotinib group, the reasons for discontinuation were AEs in two patients, withdrawal of consent in two, and the investigator's judgment of inappropriateness for study continuation in six. In the tocilizumab group, one patient discontinued due to withdrawal of consent, one due to a violation of inclusion/exclusion criteria, and five due to the investigator's judgment of inappropriateness for study continuation. After week 12, one patient in the tocilizumab group received iguratimod as rescue therapy. Eighteen patients were

enrolled before the amendment of the study protocol (ddetails are provided in Section 2.2. Changes from the previously published protocol), and eight patients were enrolled after the amendment.

Baseline characteristics were similar between the two groups, with some differences (Table 1). The median age was 74 and 65 years and the proportion of females was 76.9 and 100% in the filgotinib and tocilizumab groups, respectively. Most patients in both groups were positive for RF, and anti-CCP positivity was observed in 61.5 and 92.3% of patients in the filgotinib and tocilizumab groups, respectively. All patients in the filgotinib group and 12 patients (92.3%) in the tocilizumab group were receiving MTX at baseline. The remaining patients in the tocilizumab group, who were not receiving MTX, were treated with bucillamine 100 mg/day and tacrolimus 1 mg/day. Glucocorticoids were concomitantly administered in two patients (15.4%) in the tocilizumab group (prednisolone 5 and 3 mg/day, respectively). No patient had prior exposure to JAK inhibitors, whereas one patient in the tocilizumab group had previously received tocilizumab.

3.2. Primary endpoint

For the primary endpoint, an ACR50 response at week 12 was achieved in 38.5% (5/13) of patients in the filgotinib group and 46.2% (6/13) of those in the tocilizumab group. The risk difference was −7.69% (95% CI, −42.26–28.8), and the risk ratio was 0.83 (95% CI, 0.36–1.98) (Table 2).

3.3 Secondary endpoints

The responses for ACR20, ACR50, and ACR70 at each assessment time point are presented in Table 2. No apparent differences were observed in the achievement proportions between the filgotinib and tocilizumab groups.

Disease activity measures (CDAI, SDAI, DAS28-ESR, and DAS28-CRP) decreased from baseline to as early as week 2 in both groups (Figure 2 and Supplementary Table S3, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*). CDAI remission (CDAI ≤ 2.8) was achieved in 7.7% (1/13) of patients in the filgotinib group and 0% (0/13) in the tocilizumab group at week 2. At week 4, CDAI remission was observed in 7.7% (1/13) of patients in both groups. At week 12, CDAI remission rates were 38.5% (5/13) in the filgotinib group and 23.1% (3/13) in the tocilizumab group. At week 24, CDAI remission was achieved in 15.4% (2/13) of patients in the filgotinib group and 30.8% (4/13) in the tocilizumab group. However, no apparent between-group differences were observed in the estimated values at any time point (Supplementary Table S4, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*).

Similarly, MSUS scores (total GS score, total PD score, and GLOESS) decreased from baseline, and their median values, respectively, at week 52 were 0 in both groups, although the number of evaluable patients at this time point was limited (filgotinib group, n = 3; tocilizumab group, n = 6) (Figure 2). Notably, reductions in these scores were observed from week 4 in the filgotinib group, but from week 12 in the tocilizumab group (Supplementary Table S3, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*). However, no apparent between-group differences were observed at any time point (Supplementary Table S4, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*).

PROs (patient global VAS score, patient pain VAS

Table 2. Proportions of patients achieving ACR20/50/70 responses

	Filgotinib (n = 13)	Tocilizumab (n = 13)	Risk Difference (95% CI)	Risk Ratio (95% CI)
ACR20				
Week 2	5/13 (38.5)	6/13 (46.2)	-7.69 (-42.26, 28.80)	0.83 (0.36, 1.98)
Week 4	9/13 (69.2)	6/13 (46.2)	23.08 (-14.58, 54.95)	1.50 (0.76, 2.81)
Week 12	9/13 (69.2)	10/13 (76.9)	-7.69 (-40.37, 26.45)	0.90 (0.58, 1.42)
Week 24	8/13 (61.5)	9/13 (69.2)	-7.69 (-41.56, 27.94)	0.89 (0.52, 1.53)
Week 52	3/13 (23.1)	6/13 (46.2)	-23.08 (-54.37, 13.49)	0.50 (0.19, 1.55)
ACR50				
Week 2	2/13 (15.4)	0/13 (0.0)	15.38 (-9.27, 42.23)	Inf (0.26, 94.69)
Week 4	4/13 (30.8)	4/13 (30.8)	0.00 (-34.26, 34.26)	1.00 (0.34, 2.91)
Week 12	5/13 (38.5)	6/13 (46.2)	-7.69 (-42.26, 28.80)	0.83 (0.36, 1.98)
Week 24	7/13 (53.8)	8/13 (61.5)	-7.69 (-42.26, 28.80)	0.87 (0.47, 1.65)
Week 52	3/13 (23.1)	5/13 (38.5)	-15.38 (-47.59, 20.08)	0.60 (0.21, 1.93)
ACR70				
Week 2	0/13 (0.0)	0/13 (0.0)	0.00 (-22.81, 22.81)	0.00 (0.00, Inf)
Week 4	2/13 (15.4)	2/13 (15.4)	0.00 (-30.33, 30.33)	1.00 (0.21, 4.87)
Week 12	4/13 (30.8)	4/13 (30.8)	0.00 (-34.26, 34.26)	1.00 (0.34, 2.91)
Week 24	4/13 (30.8)	4/13 (30.8)	0.00 (-34.26, 34.26)	1.00 (0.34, 2.91)
Week 52	3/13 (23.1)	5/13 (38.5)	-15.38 (-47.59, 20.08)	0.60 (0.21, 1.93)

Values are presented as number (%). ACR, American College of Rheumatology; CI, confidence interval.

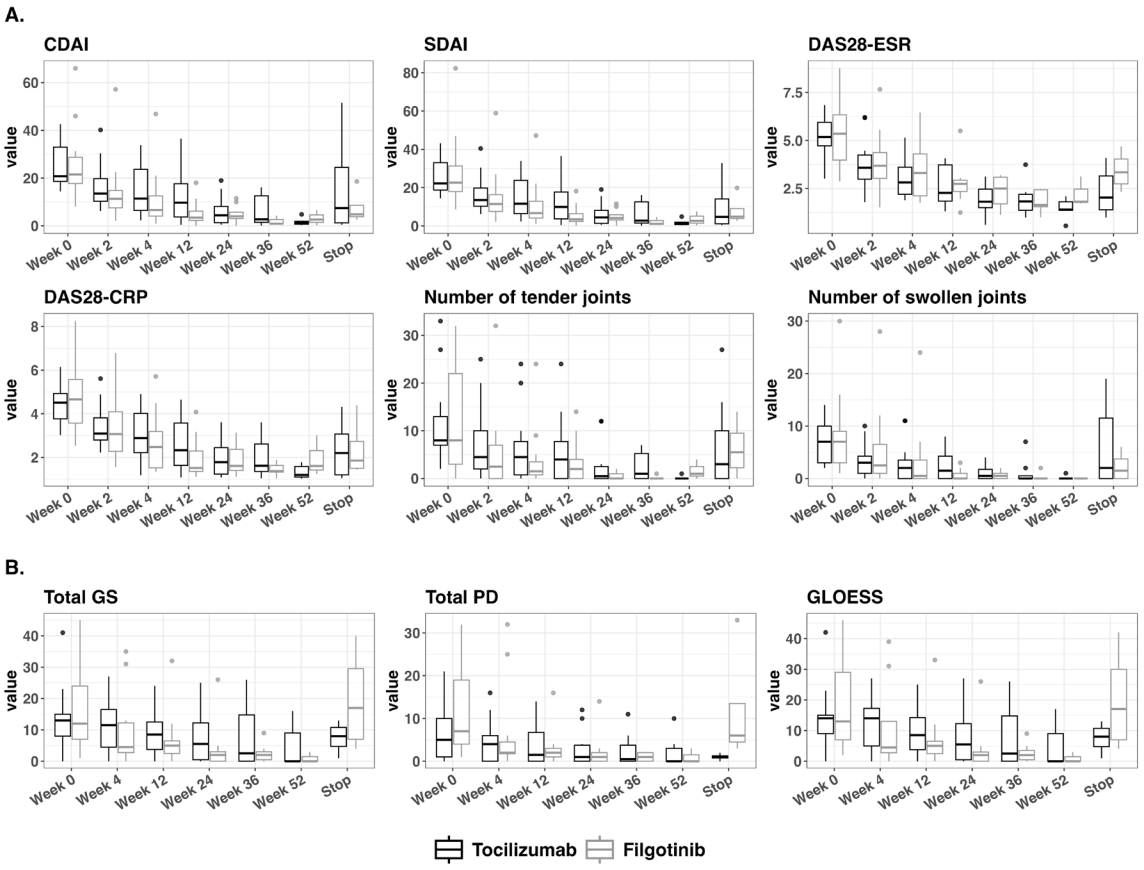

Figure 2. Values in clinical disease activity and MSUS during the study period. A. clinical disease activity and **B**. musculoskeletal ultrasound scores. Horizontal bar: median; boxes: 25th and 75th percentiles; bars: 5th and 95th percentiles. Black box/bar indicates the tocilizumab group, and gray box/bar indicates the filgotinib group. The numbers of evaluable patients at each time point were as follows. In **A**, in the tocilizumab group, the numbers were identical for CDAI, tender joint count, and swollen joint count: week 0, *n* = 13; week 2, n = 12; week 4, n = 12; week 12, n = 12; week 24, n = 10; week 36, n = 8; week 52, n = 6; and stop, n = 7. For SDAI, DAS28-ESR, and DAS28-CRP in the tocilizumab group, the numbers were week 0, n = 13; week 2, n = 12; week 4, n = 12; week 12, n = 12; week 24, n = 10; week 36, n = 8; week 52, n = 6; and stop, n = 6. In the filgotinib group, the numbers were identical across all assessments: week 0, n = 13; week 2, n = 12; week 4, n = 12; week 12, n = 11; week 24, n = 10; week 36, n = 7; week 52, n = 3; and stop, n = 4. In **B**, the numbers in the tocilizumab group were week 0, n = 13; week 4, n = 12; week 12, n = 12; week 24, n = 10; week 36, n = 8; week 52, n = 6; and stop, n = 4, whereas those in the filgotinib group were week 0, n = 13; week 4, n = 12; week 12, n = 11; week 24, n = 9; week 36, n = 7; week 52, n = 3; and stop, n = 4. Abbreviations: CDAI, Clinical Disease Activity Index; CRP, C-reactive protein; DAS28, Disease Activity Score-28; ESR, erythrocyte sedimentation rate; GLOESS, Global OMERACT-EULAR Synovitis Score; GS, grayscale; PD, power Doppler; SDAI, Simplified Disease Activity Index; MSUS, musculoskeletal ultrasound scores.

score, morning stiffness severity, duration of morning stiffness, EQ-5D-5L, FACIT-F, and HAQ-DI) also showed early improvement in both groups (Figure 3 and Supplementary Table S3, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*). At week 2, the patient global VAS scores were lower in the filgotinib group than in the tocilizumab group (− 20.96; 95% CI, − 40.16 – −1.76); however, this difference was not observed after week 4 (Table 3). In addition, other PROs showed no apparent differences between the groups at any time point (Table 3).

Changes in the vdH-mTSS were minimal, with a median change from baseline of 0 at both weeks 24 and 52 in each group (Supplementary Table S3, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*).

Supplementary Figure S2 (*https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*) shows the changes in serum cytokine and chemokine biomarkers. Serum IL-6 levels showed a sustained decrease from week 2 to week 52 in the filgotinib group, whereas the levels increased in the tocilizumab group during the same period. Moreover, from week 2, lower serum IL-6 levels were observed in the filgotinib group than in the tocilizumab group. Other than serum IL-6, no biomarkers showed consistent changes from baseline or any clear differences between the treatment groups (Supplementary Table S5, *https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*).

3.4. Subgroup analysis

The results of ACR20, ACR50, and ACR70 responses stratified by the randomization factors — disease duration of RA (< 2 years and ≥ 2 years), disease activity (DAS28-ESR > 5.1 and ≤ 5.1), and body weight (<

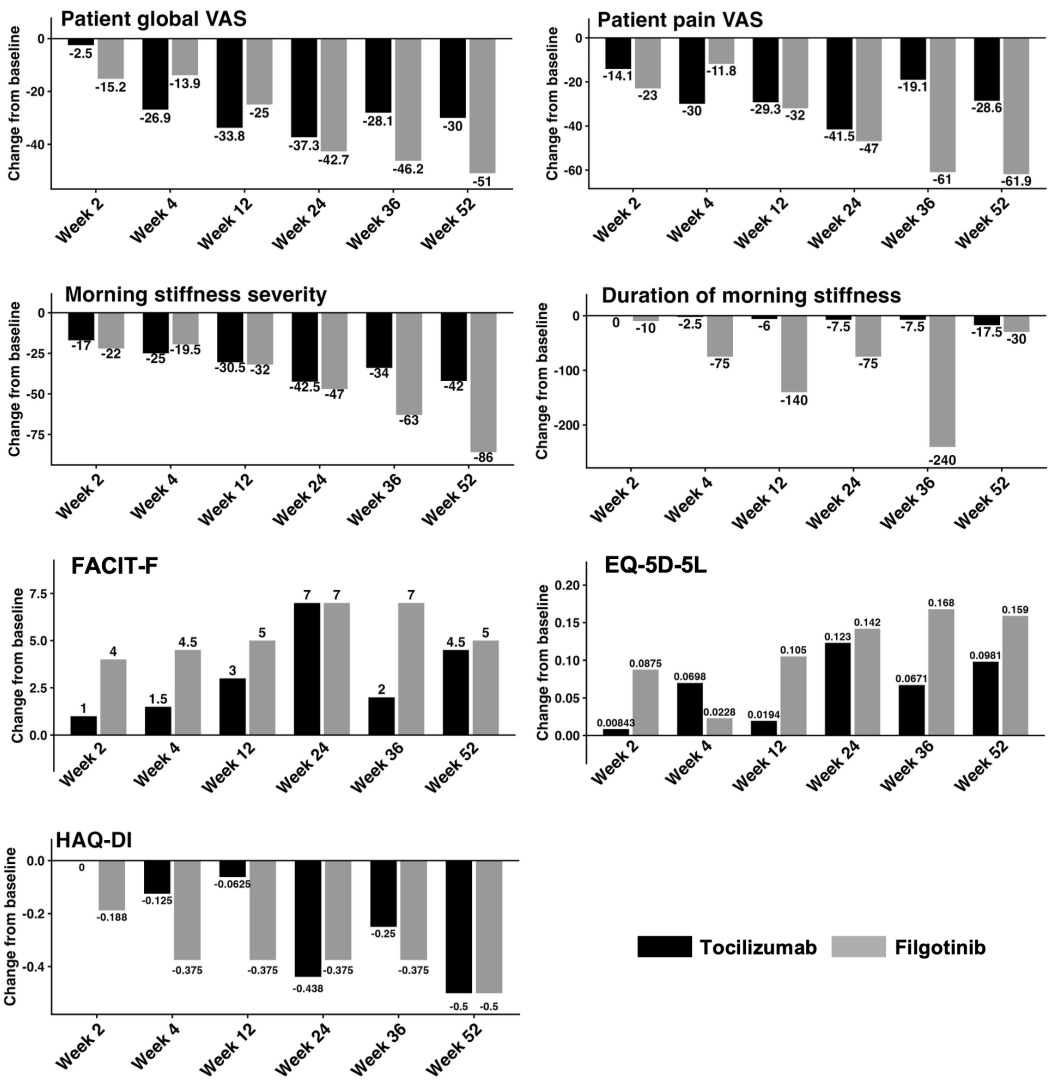

Figure 3. Changes in patient-reported outcomes during the study period. Bar graphs show the median change from baseline at each time point. The black bar indicates the tocilizumab group, and the gray bar indicates the filgotinib group. The numbers of evaluable patients at each time point were as follows. In the tocilizumab group, the numbers were identical for patient global VAS scores, patient pain VAS scores, and HAQ-DI: week 2, *n* = 12; week 4, n = 12; week 12, n = 12; week 24, n = 10; week 36, n = 8; week 52, n = 6; and stop, n = 7. For morning stiffness severity, duration of morning stiffness, FACIT-F, and EQ-5D-5L in the tocilizumab group, the numbers were week 2, n = 12; week 4, n = 12; week 12, n = 12; week 24, n = 10; week 36, n = 8; week 52, n = 6; and stop, n = 6. In the filgotinib group, the numbers were identical for patient global VAS scores, patient pain VAS scores, morning stiffness severity, duration of morning stiffness, FACIT-F, and HAQ-DI: week 2, n = 12; week 4, n = 12; week 12, n = 11; week 24, n = 10; week 36, n = 7; week 52, n = 3; and stop, n = 4. For EQ-5D-5L in the filgotinib group, the numbers were week 2, n = 11; week 4, n = 12; week 12, n = 11; week 24, n = 10; week 36, n = 7; week 52, n = 3; and stop, n = 4. Abbreviations: EQ-5D-5L, EuroQol 5 Dimension 5-Level; FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI, Health Assessment Questionnaire-Disability Index; VAS, Visual Analog Scale.

60 kg and ≥ 60 kg)—are shown in Supplementary Table S6 (*https://www.ddtjournal.com/action/getSupplementalData.php?ID=302*). No apparent between-group differences were observed for any of these subgroups.

3.5. Safety

During the 52-week treatment period, AEs were reported in three patients (23.1%) in the filgotinib group and seven patients (53.8%) in the tocilizumab group (Table 4). There were 9 AEs in the filgotinib group, including 4 serious AEs: urinary tract infection, heart failure, bacteremia, and tumor-forming pancreatitis. Treatment was discontinued in two patients in the filgotinib group due to the AEs. In the tocilizumab group, 12 AEs were reported, none of which were serious. The most common AEs were stomatitis (*n* = 4) and infections, including one case each of *Pseudomonas aeruginosa* pneumonia, periodontal disease, and upper respiratory tract infection. Treatment-related AEs occurred in 33.3% (3/9) of events in the filgotinib group and 83.3% (10/12) of events in the tocilizumab group.

Medication adherence was high in both groups, with

Table 3. Differences in patient reported outcomes between the filgotinib group and tocilizumab group

Parameter	Difference	95% CI Lower	95% CI Upper
Patient global VAS (scale: 0–100)			
Week 0	-1	-23.82	21.83
Week 2	-20.96	-40.16	-1.76
Week 4	-0.66	-21.55	20.23
Week 12	-0.49	-20.89	19.91
Week 24	3.69	-14.17	21.54
Week 36	-14.81	-35.91	6.29
Week 52	-0.99	-30.78	28.8
Patient pain VAS (scale: 0–100)			
Week 0	-3.28	-28.01	21.45
Week 2	-16.08	-37.08	4.92
Week 4	-0.26	-22.48	21.96
Week 12	-1.92	-22.21	18.38
Week 24	2.69	-12.69	18.07
Week 36	-18.34	-43.43	6.75
Week 52	1.76	-22.06	25.58
Morning stiffness severity (scale: 0–100)			
Week 0	3.38	-22.76	29.53
Week 2	-11.33	-35.26	12.59
Week 4	-3.75	-26.81	19.31
Week 12	-5.83	-25.56	13.9
Week 24	-3.6	-18.21	11.01
Week 36	-22.14	-49.57	5.28
Week 52	-1	-12.84	10.84
Duration of morning stiffness, minutes			
Week 0	275	-2.98	552.98
Week 2	37.08	-121.29	195.46
Week 4	-87.25	-349.92	175.42
Week 12	-106.01	-369.14	157.13
Week 24	124.7	-199.18	448.58
Week 36	-188.3	-610.26	233.65
Week 52	5.83	-23.86	35.53
EQ-5D-5L			
Week 0	-0.04	-0.23	0.16
Week 2	0.09	-0.13	0.3
Week 4	-0.04	-0.18	0.09
Week 12	0.06	-0.05	0.17
Week 24	-0.02	-0.16	0.11
Week 36	0.13	-0.01	0.26
Week 52	0	-0.29	0.29
FACIT-F			
Week 0	-4.62	-13.78	4.55
Week 2	-2.5	-13.60	8.6
Week 4	-1.75	-12.01	8.51
Week 12	-0.93	-9.6	7.74
Week 24	-0.9	-7.65	5.85
Week 36	6.23	-2.11	14.57
Week 52	-0.33	-23.97	23.31
HAQ-DI			
Week 0	0.16	-0.44	0.77
Week 2	-0.08	-0.73	0.57
Week 4	-0.1	-0.64	0.43
Week 12	-0.11	-0.61	0.4
Week 24	0.19	-0.48	0.85
Week 36	-0.18	-0.76	0.4
Week 52	-0.1	-1.24	1.03

Data shown the difference (Filgotinib – Tocilizumab) in each timepoint from baseline. CI, Confidence interval; EQ-5D-5L, EuroQol 5 Dimension 5-Level; FACIT-F, Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI, Health Assessment Questionnaire-Disability Index; VAS, Visual Analog Scale.

median adherence rates of 98.4% (95% CI, 94.7–99.7) in the filgotinib group and 96.3% (95% CI, 92.9–100) in the tocilizumab group.

4. Discussion

This randomized controlled trial compared the efficacy and safety of filgotinib and subcutaneous tocilizumab monotherapies in patients with RA and an inadequate response to csDMARDs. Although the planned sample size was not achieved and a formal non-inferiority

Table 4. Summary of adverse events

Event	Tocilizumab (*n* = 13)	Filgotinib (*n* = 13)
Patients with adverse event, *n* (%)[a]	7 (53.8)	3 (23.1)
Number of adverse events	12	9
Number of serious adverse events	0	4
Number of infection (serious infection)	3 (0)	3 (3)
Adverse events by severity, *n* (%)[b]		
Mild	2 (16.7)	2 (22.2)
Moderate	10 (83.3)	3 (33.3)
Severe	0 (0.0)	3 (33.3)
Life-threatening	0 (0.0)	1 (11.1)
Death	0 (0.0)	0 (0.0)
Treatment-related adverse events, *n* (%)	10 (83.3)	3 (33.3)
Patients discontinuing due to adverse events	0	2
Adverse events of special interest		
Infection	3	3
Serious infection	0	3

[a]proportion per patient, [b]proportion per event.

analysis could not be performed, the results provide meaningful insights into the comparative performance of these two therapeutic strategies. Both groups demonstrated rapid and sustained improvement in disease activity, with no apparent difference in radiographic progression between the treatments.

The introduction of bDMARDs and JAK inhibitors has greatly advanced the RA therapeutic landscape, offering multiple treatment options with distinct mechanisms of action. However, as these options have expanded, the question of how to optimally select among them has become an important clinical issue. Several studies comparing JAK or IL-6 inhibitors with TNF inhibitors show that both JAK (including filgotinib) and IL-6 inhibitors can achieve efficacy comparable to or greater than TNF inhibitors in patients with inadequate response to csDMARDs (*7-10,28-30*). Regarding comparative studies of JAK and IL-6 inhibitors, one retrospective analysis using propensity score matching reported that baricitinib and tocilizumab led to comparable improvements in DAS28-CRP, joint counts, and CRP levels at 24 weeks (*31*). Nevertheless, no randomized controlled trial has directly compared the efficacy and safety of JAK inhibitors and IL-6 inhibitors. The comparison of these two therapeutic classes holds particular scientific significance given their overlapping mechanisms of action. JAK inhibitors provide therapeutic effects through direct inhibition of the JAK-STAT signaling pathway, whereas IL-6 inhibitors indirectly suppress this pathway by blocking IL-6 signaling. Furthermore, our study addresses a clinically relevant scenario, as both filgotinib and tocilizumab demonstrate lower MTX dependency than TNF inhibitors, making monotherapy comparison particularly meaningful.

Notably, in our study, the patient global VAS score was lower in the filgotinib group than in the tocilizumab group at week 2. Recently, PROs, including the patient global VAS score, have been increasingly emphasized in RA management, as early improvements in PROs correlate with long-term pain persistence and radiographic progression (*32,33*). Previous randomized control trials have shown that baricitinib and upadacitinib lead to significantly greater and faster improvements in the patient global VAS scores than bDMARDs such as adalimumab and abatacept (*9,10,34*). A post hoc analysis of a randomized control trial showed that, among patients with an inadequate response to MTX, 200 mg filgotinib plus MTX resulted in a significantly greater improvement in the patient global VAS score at week 12 than adalimumab plus MTX (*35*). Consistent with these findings, a propensity-matched retrospective comparison between baricitinib and tocilizumab demonstrated a greater patient global VAS score improvement with baricitinib at 24 weeks (*31*). However, in our study, this early difference between the filgotinib and tocilizumab groups diminished over time, and no difference was observed in radiographic progression. Therefore, the clinical significance of early PRO improvement with filgotinib compared to tocilizumab requires further investigation.

We also evaluated MSUS findings, which are increasingly recognized as valuable tools for monitoring disease activity in RA. Previous studies demonstrate that both bDMARD and JAK inhibitor therapies improve synovial inflammation detectable by MSUS, even at early stages (*36,37*). Although clinical remission remains the primary therapeutic target to prevent joint destruction, radiographic progression occurs in 15–20% of patients with clinical remission (*38,39*). Approximately 60% of patients in clinical remission demonstrate persistent power Doppler-positive synovitis on MSUS, which represents a risk factor for subsequent joint destruction (*13,40,41*). These observations underscore the importance of performing MSUS assessment with clinical evaluation. Despite the limited sample size, both treatment groups in our study demonstrated MSUS improvement and suppression of radiographic progression as measured by the vdH-mTSS.

Multiple serum cytokines and chemokines were examined. Among all evaluated parameters, serum IL-6 was the only biomarker demonstrating consistent changes from baseline in all visits and between-group differences from week 2. Serum IL-6 levels decreased early in the filgotinib group while increasing in the tocilizumab group, with between-group differences maintained throughout the study. The elevation in serum IL-6 levels observed in the tocilizumab group is consistent with the mechanism of receptor blockade, which reduces IL-6 clearance and leads to the accumulation of circulating IL-6, correlating with receptor occupancy and antibody concentration (*42,43*). Persistently elevated serum IL-6 levels through week 52 suggest adequate antibody exposure throughout the study. In contrast, the filgotinib

group demonstrated IL-6 reduction, presumably through JAK-STAT pathway blockade. However, no definite changes were observed in other cytokines involved in the JAK–STAT pathway, including IFN-α, possibly reflecting the limited sample size and highlighting the need for further investigation.

Regarding safety, multiple AEs were observed in both groups. Although large post-marketing studies have raised concerns regarding malignancy and major adverse cardiovascular events associated with JAK inhibitors, none of these events occurred in our study (*44*). However, the limited sample size and 52-week observation period preclude conclusions regarding long-term safety. Notably, several serious AEs, including serious infections, were observed in the filgotinib group, warranting careful monitoring in clinical practice.

This study has some important limitations. First, although the study was randomized, the open-label design introduced potential treatment bias. Second, failure to achieve the predetermined target sample size precluded the planned non-inferiority analysis, limiting the analysis to descriptive comparisons. Consequently, this study does not provide definitive guidance regarding treatment selection for specific patient populations. Despite these limitations, our study provides valuable comparative data on two important therapeutic options, evaluating not only clinical activity but also MSUS findings, PROs, and biomarker profiles.

In conclusion, although the planned sample size was not attainable and robust between-group effect estimation was not feasible, both filgotinib and subcutaneous tocilizumab monotherapies led to early and sustained improvements in disease activity, including MSUS findings and PROs in patients with csDMARD-refractory RA. Although no new safety signals were identified, the occurrence of several serious AEs with filgotinib warrants careful monitoring. This trial was terminated early and represents a descriptive analysis of an underpowered randomized study. Larger, adequately powered studies are required to confirm these findings and clarify the optimal therapeutic positioning of JAK and IL-6 inhibitors in RA management.

Acknowledgements

We thank Kiyoshi Migita (Fukushima Medical University Hospital), Isao Matsumoto (University of Tsukuba Hospital), Toshihide Mimura (Saitama Medical University Hospital), Kuninobu Wakabayashi (Showa Medical University East Hospital), Shigeru Ono (Yokohama City University Medical Center), Ryusuke Yoshimi (Yokohama City University Hospital), Satoshi Ito (Niigata Prefectural Shibata Hospital), Masao Tanaka (Kyoto University Hospital), Yutaka Kawahito (Kyoto Prefectural University of Medicine Hospital), Tohru Takeuchi (Osaka Medical and Pharmaceutical University Hospital), Shiro Ohshima (Osaka Minami Medical Center), Jun Saegusa (Kobe University Hospital), Yasuhiro Tani (Nagato General Hospital), Hiroaki Dobashi (Kagawa University Hospital), Jun Ishizaki (Ehime University Hospital), Yoshinori Taniguchi (Kochi University Hospital), Yoshiya Tanaka (University of Occupational and Environmental Health Hospital), Kazuyoshi Saito (Tobata General Hospital), Tomomi Tsuru (PS Clinic), Koichiro Aratake (Ureshino Medical Center), Tamami Yoshitama (Kirishima Clinic for Rheumatic Diseases), Akitomo Okada (Nagasaki Medical Center), Norihiko Takasawa (Tohoku Medical and Pharmaceutical University Wakabayashi Hospital), Yuko Kaneko (Keio University Hospital), Hirofumi Shoda (The University of Tokyo Hospital), Masataka Kuwana (Nippon Medical School Hospital), Takehisa Ogura (Toho University Ohashi Medical Center), Sonosuke Yukawa (Yukawa Rheumatology Clinic), Takaaki Komiya (Yokohama Minami Kyosai Hospital), Norihide Hayashi (Yoshimi Hospital), Yuji Nozaki (Kindai University Hospital), Ryota Hara (Nara Medical University Hospital), Ikuko Onishi (Utazu Hospital), Yoshifumi Tada (Saga University Hospital), Tomonori Ishii (Tohoku University Hospital), Kimito Kawahata (St. Marianna University Hospital), Hiroki Ikai (Chubu Rosai Hospital), Naoki Kondo (Niigata University Medical and Dental Hospital), Shigenori Tamaki (Nagoya Rheumatology Clinic), Shinichi Mizuki (Matsuyama Red Cross Hospital), Takuya Nakazawa (Chiba-East-Hospital), and Tsuyoshi Watanabe (National Center for Geriatrics Gerontology), all of whom are investigators of joint research institutions. We thank our colleagues and staff at the Department of Immunology and Rheumatology, Nagasaki University Hospital, for their support. We thank Editage (www.editage.jp) for English language editing.

Funding: This study was funded by Gilead Sciences, Inc. (Foster City, CA, USA).

Conflict of Interest: TS received speaker's fees from AbbVie GK, Asahi Kasei Pharma Corporation, Astellas Pharma Inc., Chugai Pharmaceutical Co., and Eisai Co. SK received speaker's fees from AbbVie GK. TK received speaker's fees from AbbVie GK, Asahi Kasei Pharma Corporation, Boehringer Ingelheim Japan, Chugai Pharmaceutical Co., Eisai Co., Eli Lilly Japan, Gilead Sciences, Inc, and Taisho Pharmaceutical Co. SS received speaker's fees from Pfizer Inc. KM received grant support from ONO Pharmaceutical Co., and speaker's fees and/or honoraria from AbbVie GK, Asahi Kasei Pharma Corp., Astellas Pharma Inc., Boehringer Ingelheim Japan, Chugai Pharmaceutical Co., Eisai Co., Eli Lilly Japan, Gilead Sciences, Inc., Kyowa Kirin Co., Ltd., Novartis AG., Mitsubishi Tanabe Pharma Co., Mochida Pharmaceutical Co., Ltd., ONO Pharmaceutical Co., Pfizer Inc., Taisho Pharmaceutical Co., Takeda Pharmaceutical Company Limited, UCB Japan Co.,

Ltd., and Viatris Inc. SH received speaker's fees from AbbVie GK, Asahi Kasei Pharma Corporation, Chugai Pharmaceutical Co., Eisai Co., Eli Lilly Japan, Taisho Pharmaceutical Co., UCB Japan Co. DN received grant support from Asahi Kasei Pharma Corp., and Chugai Pharmaceutical Co., and speaker's fees from AbbVie GK, Asahi Kasei Pharma Corp., Chugai Pharmaceutical Co., and Taisho Pharmaceutical Co. MO received speaker's fees from Pfizer Inc., Eli Lilly Japan K.K., Teijin Healthcare Limited, Mitsubishi Tanabe Pharma Corporation, Daiichi Sankyo Company, Limited, Astellas Pharma Inc., AbbVie Inc., Bristol Myers Squibb K.K., Celltrion Healthcare Japan Co., Ltd., Taisho Pharmaceutical Co., Ltd., Eisai Co., Ltd., Eisai Korea Inc., Takeda Pharmaceutical Company Limited, and Chugai Pharmaceutical Co., Ltd. NT received research grants from Asahi Kasei Pharma Corp., Ayumi Pharmaceutical Corp., Bristol Myers Squibb Co., Ltd., Chugai Pharmaceutical Co., Ltd., Nippon Boehringer Ingelheim Co., Ltd., Taisho Pharmaceutical Co., Ltd., and speaker's fees from AstraZeneca plc, AbbVie Inc., Eli Lilly Japan K.K., GlaxoSmithKline K.K., Kissei Pharmaceutical Co., Ltd., Otsuka Pharmaceutical Co., Ltd., and UCB Japan Co., Ltd. HK received speaker's fees from AbbVie GK, Asahi Kasei Pharma Corporation, and Astellas Pharma Inc. TA received grant support from GlaxoSmithKline plc., Bristol-Myers Squibb Co., Zenyaku Kogyo Company, Ltd., Amgen K.K., KISSEI PHARMACEUTICAL CO., LTD., AbbVie Inc., Mitsubishi Tanabe Pharma, Eisai Co. Ltd., Nippon Boehringer Ingelheim Co., Otsuka Pharmaceutical Co., Ltd., and Chugai Pharmaceutical Co., Ltd., consulting fees from AstraZeneca plc., Idorsia Pharmaceuticals Ltd., Otsuka Pharmaceutical Co., Gilead Sciences, Inc., GlaxoSmithKline plc., Sanofi K.K., Eli Lilly Japan K.K., Nippon Boehringer Ingelheim Co., Ltd., Janssen Pharmaceutical K.K., UCB Japan Co. Ltd., KISSEI PHARMACEUTICAL CO., LTD., and Novartis Pharma K.K., and speaker's fees from Bristol-Myers Squibb Co., Eli Lilly Japan K.K., Nippon Boehringer Ingelheim Co., Ltd., Eisai Co. Ltd., AbbVie Inc., Pfizer Inc., and Gilead Sciences, Inc. TO received speaker's fees and/or research grants from AbbVie GK, Asahi Kasei Pharma Corporation, Astellas Pharma Inc., Daiichi Sankyo Co., Ltd., Eli Lilly Japan, and UCB Japan Co., Ltd. HT received speaker's fees from Gilead Sciences, Inc. TH received speaker's fees from AbbVie GK, Eli Lilly Japan, Pfizer Inc., Asahi Kasei Pharma Corp., Bristol-Myers Squibb Company, Chugai Pharmaceutical Co., Janssen Pharmaceutical K.K., Taisho Pharmaceutical Co., and Eisai Co. SF received speaker's fees from Daiichi Sankyo Co., Ltd., Kissei Pharmaceutical Co., Ltd., AstraZeneca PLC, Pfizer Inc., Torii Pharmaceutical Co., Ltd., Boehringer Ingelheim Japan, and Novartis Pharmaceuticals, and consulting fees from AstraZeneca PLC, Novartis Pharmaceuticals, and Asahi Kasei Pharma Corp. AK received grant support from AbbVie GK, Argenx SE, Asahi Kasei Pharma Corp., Astellas Pharma Inc., Ayumi Pharmaceutical Corp., Boehringer Ingelheim Japan, Celltrion Healthcare Co., Chugai Pharmaceutical Co., Gilead Sciences, Inc., ONO Pharmaceutical Co., Neopharma LLC., and Taisho Pharmaceutical Co., and speaker's fees from AbbVie GK, Actelion Pharmaceuticals Ltd., Asahi Kasei Pharma Co., Astellas Pharma Inc., Boehringer Ingelheim Japan, Chugai Pharmaceutical Co., Daiichi Sankyo Co., Ltd., Eisai Co., Eli Lilly Japan, Kissei Pharmaceutical Co., Ltd., Mitsubishi Tanabe Pharma Co., ONO Pharmaceutical Co., Pfizer Inc., Taisho Pharmaceutical Co., and UCB Japan Co., Ltd. The other authors have no conflicts of interest to declare.

This was an investigator-initiated study. The study protocol was designed by the authors and reviewed by Gilead Sciences, Inc. for feasibility before funding approval. Gilead Sciences, Inc. provided financial support for the study but had no role in the study design; data collection; data analysis or interpretation; or the decision to submit the manuscript for publication. The sponsor was given the opportunity to review the manuscript; no comments were provided. The authors had full access to all data and take final responsibility for the content of the manuscript.

Institutional Review Board Statement: The study protocol was approved by the certified review board of Nagasaki University (Reference number: CRB20-026).

Informed Consent Statement: Written informed consent was provided by the patients before enrollment.

Data Availability Statement: The datasets used or analyzed (or both) during the current study are available from the corresponding author upon reasonable request.

References

1. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. Lancet. 2010; 376:1094-1108.
2. Smolen JS, Breedveld FC, Burmester GR, *et al*. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis. 2016; 75:3-15.
3. Smolen JS, Landewé RBM, Bergstra SA, *et al.* EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2022 update. Ann Rheum Dis. 2022; 82:3-18.
4. Harigai M, Kaneko Y, Tanaka E, *et al*. 2024 Update of the Japan College of Rheumatology Clinical Practice Guidelines for the Management of Rheumatoid Arthritis: Secondary publication. Mod Rheumatol. 2025; 35:387-401.
5. Gadina M, Le MT, Schwartz DM, Silvennoinen O, Nakayamada S, Yamaoka K, O'Shea JJ. Janus kinases to jakinibs: from basic insights to clinical practice. Rheumatology (Oxford). 2019; 58:i4-i16.

6. Genovese MC, Kalunian K, Gottenberg JE, Mozaffarian N, Bartok B, Matzkies F, Gao J, Guo Y, Tasset C, Sundy JS, de Vlam K, Walker D, Takeuchi T. Effect of Filgotinib vs Placebo on Clinical Response in Patients With Moderate to Severe Rheumatoid Arthritis Refractory to Disease-Modifying Antirheumatic Drug Therapy: The FINCH 2 Randomized Clinical Trial. JAMA. 2019; 322:315-325.
7. Combe B, Kivitz A, Tanaka Y, *et al*. Filgotinib versus placebo or adalimumab in patients with rheumatoid arthritis and inadequate response to methotrexate: a phase III randomised clinical trial. Ann Rheum Dis. 2021; 80:848-858.
8. Fleischmann R, Pangan AL, Song IH, Mysler E, Bessette L, Peterfy C, Durez P, Ostor AJ, Li Y, Zhou Y, Othman AA, Genovese MC. Upadacitinib Versus Placebo or Adalimumab in Patients With Rheumatoid Arthritis and an Inadequate Response to Methotrexate: Results of a Phase III, Double-Blind, Randomized Controlled Trial. Arthritis Rheumatol. 2019; 71:1788-1800.
9. Fleischmann RM, Genovese MC, Enejosa JV, Mysler E, Bessette L, Peterfy C, Durez P, Ostor A, Li Y, Song IH. Safety and effectiveness of upadacitinib or adalimumab plus methotrexate in patients with rheumatoid arthritis over 48 weeks with switch to alternate therapy in patients with insufficient response. Ann Rheum Dis. 2019; 78:1454-1462.
10. Taylor PC, Keystone EC, van der Heijde D, *et al*. Baricitinib versus Placebo or Adalimumab in Rheumatoid Arthritis. N Engl J Med. 2017; 376:652-662.
11. Nguyen H, Ruyssen-Witrand A, Gandjbakhch F, Constantin A, Foltz V, Cantagrel A. Prevalence of ultrasound-detected residual synovitis and risk of relapse and structural progression in rheumatoid arthritis patients in clinical remission: a systematic review and meta-analysis. Rheumatology (Oxford). 2014; 53:2110-2118.
12. Kawashiri SY, Suzuki T, Nakashima Y, Horai Y, Okada A, Iwamoto N, Ichinose K, Tamai M, Arima K, Nakamura H, Origuchi T, Uetani M, Aoyagi K, Eguchi K, Kawakami A. Ultrasonographic examination of rheumatoid arthritis patients who are free of physical synovitis: power Doppler subclinical synovitis is associated with bone erosion. Rheumatology (Oxford). 2014; 53:562-569.
13. Han J, Geng Y, Deng X, Zhang Z. Subclinical Synovitis Assessed by Ultrasound Predicts Flare and Progressive Bone Erosion in Rheumatoid Arthritis Patients with Clinical Remission: A Systematic Review and Metaanalysis. J Rheumatol. 2016; 43:2010-2018.
14. Zufferey P, Scherer A, Nissen MJ, Ciurea A, Tamborrini G, Brulhart L, Blumhardt S, Toniolo M, Möller B, Ziswiler HR. Can Ultrasound Be Used to Predict Loss of Remission in Patients with RA in a Real-life Setting? A Multicenter Cohort Study. J Rheumatol. 2018; 45:887-894.
15. Colebatch AN, Edwards CJ, Østergaard M, *et al*. EULAR recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis. Ann Rheum Dis. 2013; 72:804-814.
16. D'Agostino MA, Terslev L, Wakefield R, Østergaard M, Balint P, Naredo E, Iagnocco A, Backhaus M, Grassi W, Emery P. Novel algorithms for the pragmatic use of ultrasound in the management of patients with rheumatoid arthritis: from diagnosis to remission. Ann Rheum Dis. 2016; 75:1902-1908.
17. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013; 310:2191-2194.
18. Shimizu T, Kawashiri SY, Morimoto S, Kawazoe Y, Kuroda S, Kawasaki R, Ito Y, Kiya R, Sato S, Yamamoto H, Kawakami A. Efficacy and safety of selective JAK 1 inhibitor filgotinib in active rheumatoid arthritis patients with inadequate response to methotrexate: comparative study with filgotinib and tocilizumab examined by clinical index as well as musculoskeletal ultrasound assessment (TRANSFORM study): study protocol for a randomized, open-label, parallel-group, multicenter, and non-inferiority clinical trial. Trials. 2023; 24:161.
19. Aletaha D, Neogi T, Silman AJ, *et al*. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010; 62:2569-2581.
20. Prevoo ML, van 't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum. 1995; 38:44-48.
21. Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. Arthritis Rheum. 1980; 23:137-145.
22. Torp-Pedersen ST, Terslev L. Settings and artefacts relevant in colour/power Doppler ultrasound in rheumatology. Ann Rheum Dis. 2008; 67:143-149.
23. D'Agostino MA, Boers M, Wakefield RJ, Berner Hammer H, Vittecoq O, Filippou G, Balint P, Möller I, Iagnocco A, Naredo E, Østergaard M, Gaillez C, Le Bars M. Exploring a new ultrasound score as a clinical predictive tool in patients with rheumatoid arthritis starting abatacept: Results from the APPRAISE study. RMD Open. 2016; 2:e000237.
24. Tanaka Y, Oba K, Koike T, *et al.* Sustained discontinuation of infliximab with a raising-dose strategy after obtaining remission in patients with rheumatoid arthritis: the RRRR study, a randomised controlled trial. Ann Rheum Dis. 2020; 79:94-102.
25. van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. J Rheumatol. 2000; 27:261-263.
26. Nishimoto N, Yoshizaki K, Miyasaka N, Yamamoto K, Kawai S, Takeuchi T, Hashimoto J, Azuma J, Kishimoto T. Treatment of rheumatoid arthritis with humanized anti-interleukin-6 receptor antibody: a multicenter, double-blind, placebo-controlled trial. Arthritis Rheum. 2004; 50:1761-1769.
27. Mee RW. Confidence bounds for the difference between two probabilities. Biometrics. 1984; 40:1175-1176.
28. Fleischmann R, Mysler E, Hall S, *et al*. Efficacy and safety of tofacitinib monotherapy, tofacitinib with methotrexate, and adalimumab with methotrexate in patients with rheumatoid arthritis (ORAL Strategy): a phase 3b/4, double-blind, head-to-head, randomised controlled trial. Lancet. 2017; 390:457-468.
29. Burmester GR, Lin Y, Patel R, van Adelsberg J, Mangan EK, Graham NM, van Hoogstraten H, Bauer D, Ignacio Vargas J, Lee EB. Efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy for the treatment of patients with active rheumatoid arthritis (MONARCH): a randomised, double-blind, parallel-group phase III trial. Ann Rheum Dis. 2017; 76:840-847.
30. Gabay C, Emery P, van Vollenhoven R, Dikranian A,

Alten R, Pavelka K, Klearman M, Musselman D, Agarwal S, Green J, Kavanaugh A. Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (ADACTA): a randomised, double-blind, controlled phase 4 trial. Lancet. 2013; 381:1541-1550.

31. Asai S, Takahashi N, Kobayakawa T, *et al*. Comparison of the effects of baricitinib and tocilizumab on disease activity in patients with rheumatoid arthritis: a propensity score matching analysis. Clin Rheumatol. 2021; 40:3143-3151.
32. Strand V, Lee EB, Yazici Y, Dikranian A, Wilkinson B, Takiya L, Zang C, Bananis E, Bergman MJ. Evaluation of disease activity in patients with rheumatoid arthritis treated with tofacitinib by RAPID3: post hoc analyses from two phase 3 trials. Clin Rheumatol. 2018; 37:2043-2053.
33. Eberhard A, Bergman S, Mandl T, Olofsson T, Sharma A, Turesson C. Joint tenderness at 3 months follow-up better predicts long-term pain than baseline characteristics in early rheumatoid arthritis patients. Rheumatology (Oxford). 2024; 63:734-741.
34. Rubbert-Roth A, Enejosa J, Pangan AL, Haraoui B, Rischmueller M, Khan N, Zhang Y, Martin N, Xavier RM. Trial of Upadacitinib or Abatacept in Rheumatoid Arthritis. N Engl J Med. 2020; 383:1511-1521.
35. Bingham CO, 3rd, Walker D, Nash P, Lee SJ, Ye L, Hu H, Khalid JM, Combe B. The impact of filgotinib on patient-reported outcomes and health-related quality of life for patients with active rheumatoid arthritis: a post hoc analysis of Phase 3 studies. Arthritis Res Ther. 2022; 24:11.
36. Nishino A, Kawashiri SY, Koga T, *et al*. Ultrasonographic Efficacy of Biologic and Targeted Synthetic Disease-Modifying Antirheumatic Drug Therapy in Rheumatoid Arthritis From a Multicenter Rheumatoid Arthritis Ultrasound Prospective Cohort in Japan. Arthritis Care Res (Hoboken). 2018; 70:1719-1726.
37. Endo Y, Kawashiri SY, Nishino A, *et al*. Ultrasound efficacy of targeted-synthetic disease-modifying anti-rheumatic drug treatment in rheumatoid arthritis: a multicenter prospective cohort study in Japan. Scand J Rheumatol. 2022; 51:259-267.
38. Brown AK, Conaghan PG, Karim Z, Quinn MA, Ikeda K, Peterfy CG, Hensor E, Wakefield RJ, O'Connor PJ, Emery P. An explanation for the apparent dissociation between clinical remission and continued structural deterioration in rheumatoid arthritis. Arthritis Rheum. 2008; 58:2958-2967.
39. Raffeiner B, Grisan E, Botsios C, Stramare R, Rizzo G, Bernardi L, Punzi L, Ometto F, Doria A. Grade and location of power Doppler are predictive of damage progression in rheumatoid arthritis patients in clinical remission by anti-tumour necrosis factor α. Rheumatology (Oxford). 2017; 56:1320-1325.
40. Brown AK, Quinn MA, Karim Z, Conaghan PG, Peterfy CG, Hensor E, Wakefield RJ, O'Connor PJ, Emery P. Presence of significant synovitis in rheumatoid arthritis patients with disease-modifying antirheumatic drug-induced clinical remission: evidence from an imaging study may explain structural progression. Arthritis Rheum. 2006; 54:3761-3773.
41. Ten Cate DF, Luime JJ, Swen N, Gerards AH, De Jager MH, Basoski NM, Hazes JM, Haagsma CJ, Jacobs JW. Role of ultrasonography in diagnosing early rheumatoid arthritis and remission of rheumatoid arthritis--a systematic review of the literature. Arthritis Res Ther. 2013; 15:R4.
42. Nishimoto N, Terao K, Mima T, Nakahara H, Takagi N, Kakehi T. Mechanisms and pathologic significances in increase in serum interleukin-6 (IL-6) and soluble IL-6 receptor after administration of an anti-IL-6 receptor antibody, tocilizumab, in patients with rheumatoid arthritis and Castleman disease. Blood. 2008; 112:3959-3964.
43. Zhang X, Georgy A, Rowell L. Pharmacokinetics and pharmacodynamics of tocilizumab, a humanized anti-interleukin-6 receptor monoclonal antibody, following single-dose administration by subcutaneous and intravenous routes to healthy subjects. Int J Clin Pharmacol Ther. 2013; 51:443-455.
44. Ytterberg SR, Bhatt DL, Mikuls TR, Koch GG, Fleischmann R, Rivas JL, Germino R, Menon S, Sun Y, Wang C, Shapiro AB, Kanik KS, Connell CA. Cardiovascular and Cancer Risk with Tofacitinib in Rheumatoid Arthritis. N Engl J Med. 2022; 386:316-326.

Received February 15, 2026; Revised April 20, 2026; Accepted April 23, 2026.

**Address correspondence to:*
Dr. Shin-ya Kawashiri, Department of Community Medicine, Division of Advanced Preventive Medical Sciences, Nagasaki University Graduate School of Biomedical Sciences, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan.
E-mail: shin-ya@nagasaki-u.ac.jp

Released online in J-STAGE as advance publication April 27, 2026.

Brief Report

DOI: 10.5582/ddt.2026.01000

Effects of Janus kinase inhibition and interleukin 6 inhibition on serum cytokine/chemokine in idiopathic multicentric Castleman disease

Shoichi Fukui[1,2], Remi Sumiyoshi[1,2], Tomohiro Koga[1], Naoki Hosogaya[2], Sawana Narita[2], Shimpei Morimoto[2], Osamu Kamisawa[2], Rieko Kiya[2], Atsushi Katsube[3], Shingo Yano[3], Atsushi Kawakami[1,*]

[1] Department of Immunology and Rheumatology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan;
[2] Clinical Research Center, Nagasaki University Hospital, Nagasaki, Japan;
[3] Division of Clinical Oncology/Hematology, Department of Internal Medicine, The Jikei University School of Medicine, Tokyo, Japan.

SUMMARY: Idiopathic multicentric Castleman disease (iMCD) is a rare lymphoproliferative disease characterized by systemic inflammation. Although IL-6 receptor blockade with tocilizumab is an established treatment, Janus kinase (JAK) inhibition may offer broader immunomodulation by targeting multiple cytokine signaling pathways. This comparative longitudinal study evaluated 41 serum cytokines/chemokines in 10 patients with plasma-cell-type iMCD (five treated with filgotinib (JAK1 preferential inhibitor) for 52 weeks and five treated with tocilizumab for a median of 28 months) to compare the cytokine/chemokine suppression profiles between these two mechanistically distinct therapies. Patient values were normalized to those of healthy controls (n = 101) using Z-scores. At baseline, both groups exhibited marked elevations in cytokines, including IL-12p70, IL-22, IFN-γ, and IL-6. Both treatments resulted in significant changes in multiple cytokines, with 12 cytokines showing significant changes in each group. Between-group comparison revealed only three cytokines with differential responses: IL-6 (receptor blockade artifact), IL-15 (greater suppression with filgotinib), and PDGF-AA (greater suppression with tocilizumab). Overall, cytokine suppression was equivalent between the treatments (median ΔZ-score: −0.32 (filgotinib) *vs.* −0.27 (tocilizumab), p = 0.74). Principal component analysis demonstrated parallel treatment trajectories toward normalization. These exploratory findings from this small-sample study suggest that JAK inhibition and IL-6 receptor blockade may achieve comparable broad-spectrum cytokine/chemokine suppression in iMCD, despite the different mechanisms of action. However, the discordance between cytokine suppression and clinical improvement with filgotinib suggests that complete IL-6 pathway blockade remains critical for the clinical response in iMCD, highlighting the need to identify additional pathogenic drivers beyond the measured cytokine network.

Keywords: multicentric Castleman disease, cytokine profile, filgotinib, tocilizumab, JAK inhibitor, biomarker

1. Introduction

Idiopathic multicentric Castleman disease (iMCD) is a rare lymphoproliferative disease characterized by polyclonal lymph node hyperplasia, distinctive histopathological features, and systemic inflammatory symptoms (*1,2*). iMCD is a rare and refractory disease that annually affects ~900-4,200 people in the US (*3*) and ~1,500 in Japan (*4*). Patients typically present with generalized lymphadenopathy accompanied by constitutional symptoms, including fever, fatigue, weight loss, and night sweats. Laboratory abnormalities are hallmarks of the disease and include elevated C-reactive protein (CRP) levels, anemia of chronic inflammation, hypoalbuminemia, and elevated erythrocyte sedimentation rate. The constellation of clinical and laboratory findings reflects the systemic inflammatory state characteristic of iMCD. Without effective treatment, iMCD can progress to life-threatening complications, such as multiorgan damage and secondary amyloidosis.

The pathogenesis of iMCD remains incompletely understood; however, interleukin-6 (IL-6) has been identified as a central driver of disease manifestations through extensive research. Elevated serum IL-6 levels have been reported in patients with Castleman disease (*5*). In addition, single-cell RNA sequencing and spatial enhanced resolution omics sequencing demonstrated that IL-6 pathway signals were dominant in nodal fibroblastic

reticular cells and endothelial cells in the lymph nodes of iMCD (*6*). Anti-IL-6 therapies, including the IL-6 receptor antibody tocilizumab (approved in Japan) (*7*) and the IL-6-neutralizing antibody siltuximab (approved in the United States, Europe, and other regions) (*8*), have transformed the patient outcomes. However, a significant proportion of patients do not benefit from anti–IL-6 treatment, and additional therapeutic options are needed for non-responders, especially severely afflicted patients (*9*).

The inflammatory milieu of iMCD extends beyond IL-6, encompassing a complex network of a pleomorphic cytokine profile, and the disease is not driven by IL-6. Serum proteomic analyses have revealed a "chemokine storm" in iMCD, with the simultaneous elevation of numerous inflammatory mediators that may perpetuate disease activity (*10*). This observation has prompted interest in therapeutic strategies that target broader inflammatory networks rather than individual cytokines.

To regulate broader inflammatory networks, Janus kinases (JAKs), which are downstream of cytokine receptors, including IL-6 (*11*), appear to be promising treatment targets for iMCD. JAKs contribute to the signal transducer and activator of transcription (STAT) 3 activation in iMCD pathogenesis (*12*), suggesting that JAK-STAT signaling inhibition is a potential treatment approach. This rationale led to a phase Ib investigator-led clinical trial of filgotinib, a JAK1 preferential inhibitor, for iMCD, which was conducted across Japan in 2024 (*13*).

Filgotinib preferentially inhibits JAK1-dependent cytokine signaling *in vitro* (*14*) and has been approved for use in rheumatoid arthritis (RA) (*15*) and ulcerative colitis (*16*). JAK1 inhibition theoretically offers broader immunomodulation by simultaneously suppressing the signaling of multiple cytokine receptors that utilize JAK1 for signal transduction. This mechanistic difference raises the fundamental question of whether JAK1 inhibition might achieve more comprehensive cytokine suppression than selective IL-6 receptor blockade, potentially offering advantages for patients with complex cytokine dysregulation.

In an evaluation at eight weeks, treatment with filgotinib did not show apparent efficacy on the CHAP score and its components (CRP, hemoglobin, and albumin) (*13*). Evaluation of multiple cytokines and chemokines demonstrated differences in the serum levels of FGF-2, IL-4, IL-6, TNF-β, and VEGF-A at eight weeks between filgotinib- and tocilizumab-treated patients (*17*). However, the evaluation at eight weeks seemed too early to draw a definitive conclusion.

This comparative longitudinal study aimed to comprehensively characterize the cytokine and chemokine profiles of both treatments, with normalization to healthy controls. Our approach enables a direct comparison of different cytokines and chemokines on a common scale using Z-score normalization, facilitating the interpretation of the relative magnitude of abnormalities and treatment effects. Specifically, we sought to (1) compare the breadth and magnitude of cytokine/chemokine suppression between filgotinib and tocilizumab across the entire inflammatory network; (2) identify cytokine/chemokine biomarkers associated with clinical parameters and treatment response; (3) determine whether JAK1 inhibition achieves broader immunomodulation than IL-6 receptor blockade; and (4) characterize the specific cytokines that show differential responses to these two mechanistically distinct therapies. Understanding these cytokine/chemokine dynamics has important implications for future therapeutic strategies for this challenging rare disease.

2. Materials and Methods

This comparative longitudinal study evaluated serum cytokine and chemokine dynamics in two cohorts of patients with iMCD diagnosed according to the Japanese Ministry of Health, Labour and Welfare criteria (designated intractable diseases notice no. 331, effective April 1, 2018). The filgotinib cohort comprised 5 patients enrolled in a prospective, single-arm, open-label Phase Ib clinical trial (registered in the Japan Registry of Clinical Trials (https://jrct.niph.go.jp/) as jRCT2071230108 approved by the Nagasaki University Hospital Institutional Review Board, approval No. 123-002 and jRCTs071230120 approved by the Nagasaki University Hospital Institutional Review Board, approval No. CRB23-008). Patients with a total score on the CHAP (C-reactive protein (CRP), Hemoglobin, Albumin, + Performance Status (PS, Eastern Cooperative Oncology Group [ECOG])) (*4*) that was ≥ 2 points in total with hemoglobin or albumin ≥ 1 point and CRP ≥ 1 point at baseline were included. Patients with iMCD-TAFRO (thrombocytopenia, anasarca, fever, reticulin fibrosis, renal insufficiency, and organomegaly clinical subtype), as defined by the validated international definition (*18*), were excluded. These patients received filgotinib (Gilead Sciences, Inc., Foster City, CA, USA) 200 mg orally once daily for 52 weeks, and serum samples were collected. The tocilizumab cohort comprised five patients receiving standard tocilizumab (Chugai Pharmaceutical Co., Ltd., Tokyo, Japan) therapy according to the approved dosing regimen, with serum samples collected at baseline and post-treatment status (median 28 months [minimum 11 - maximum 73 months]). All patients had plasma-cell-type histological features on lymph node biopsy and were human herpesvirus 8 (HHV-8) negative. The baseline characteristics are summarized in Table 1.

This study complied with the Declaration of Helsinki and was approved by the Nagasaki University Hospital Institutional Review Board (approval No. 25111306). An opt-out approach was used to obtain patient consent for the study.

The sera of the patients were analyzed using the

Table 1. Characteristics of patients with idiopathic multicentric Castleman disease treated with either tocilizumab or with filgotinib at baseline and after treatment

Characteristics	Tocilizumab (*n* = 5)	Filgotinib (*n* = 5)
At baseline		
Age, years, median (min−max)	62 (43−69)	60 (37−61)
Sex, female, *n* (%)	2 (40)	2 (40)
Height, cm, median (min−max)	160 (152−172)	170 (152−176)
Body weight, kg, median (min−max)	55.7 (46.1−72.5)	73.7 (50.9−75.0)
Treatment-naïve, *n* (%)	5 (100)	3 (60)
Previous immunosuppressant, *n* (%)		
Prednisolone	0 (0)	2 (40)
Others	0 (0)	0 (0)
Concomitant immunosuppressant, *n* (%)		
Prednisolone 10 mg/day	0 (0)	1 (20)
Others	0 (0)	0 (0)
Histology, *n* (%)		
Hyaline vascular type	0 (0)	0 (0)
Plasma cell type	5 (100)	5 (100)
Mixed type	0 (0)	0 (0)
CRP, mg/dL, median (min−max)	4.90 (3.70−7.22)	6.73 (2.47−7.45)
Hemoglobin, g/dL, median (min−max)	9.3 (8.5−11.6)	10.1 (8.1−11.8)
Albumin, g/dL, median (min−max)	2.4 (2.4−2.7)	2.9 (2.5−3.3)
ECOG-PS, *n* (%)		
0	4 (80)	2 (40)
1	1 (20)	3 (60)
CHAP score, median (min−max)	5 (3−7)	5 (3−6)
Platelet count, $\times 10^3/\mu L$, median (min−max)	338 (262−456)	339 (226−556)
Immunoglobulin G, mg/dL, median (min−max)	4,436 (3,733−4,996)	5,357 (4,446−5,968)
At 52 weeks* or later**		
CRP, mg/dL, median (min−max)	0.01 (0.01−0.10)	3.30 (1.92−6.31)
Hemoglobin, g/dL, median (min−max)	14.1 (12.6−14.3)	11.1 (9.4−12.3)
Albumin, g/dL, median (min−max)	4.2 (4.0−4.3)	3.3 (2.8−3.6)
ECOG-PS, *n* (%)		
0	5 (100)	4 (80)
1	0 (0)	1 (20)
CHAP score, median (min−max)	0 (0−0)	3 (1−4)

Abbreviation: CHAP: CRP, hemoglobin, albumin, + PS (ECOG), CRP: C-reactive protein, ECOG-PS: Eastern Cooperative Oncology Group-Performance Status. * for the filgotinib group, ** for the tocilizumab group.

Milliplex® MAP Human Cytokine/Chemokine Magnetic Bead Panel-Premixed 41 Plex panel (Merck Millipore, Billerica, MA, USA) and MAGPIX® with xPONENT® software (Luminex Corp., Austin, TX, USA). The levels of 41 cytokines/chemokines were measured as follows: epidermal growth factor (EGF), CCL11/eotaxin, basic fibroblast growth factor (FGF-2/bFGF), FMS-like tyrosine kinase 3 ligand (FLT-3 L), fractalkine, granulocyte colony-stimulating factor (G-CSF), granulocyte-macrophage colony-stimulating factor (GM-CSF), chemokine (C-X-C motif) ligand 1 (CXCL1/GRO-α), interferon (IFN)-α2, IFN-γ, interleukin (IL)-10, IL-12p40, IL-12p70, IL-13, IL-15, IL-17A, IL-17F, IL-18, IL-1 receptor antagonist (IL-1RA), IL-1α, IL-1β, IL-2, IL-22, IL-27, IL-4, IL-5, IL-6, IL-7, IL-8, C-X-C motif chemokine ligand 10 (CXCL10/IP-10), monocyte chemoattractant protein-1 (MCP-1/CCL2), MCP-3, macrophage-derived chemokine (MDC), CCL3/macrophage inflammatory protein (MIP)-1α, CCL4/MIP-1β, platelet-derived growth factor (PDGF)-AA, transforming growth factor (TGF)-α, tumor necrosis factor (TNF)-α, TNF-β, vascular endothelial growth factor (VEGF)-A, and soluble CD40 ligand (sCD40L).

Residents of the town of Saza in Nagasaki prefecture who underwent specific health checkups in 2016 were used for the 41 Plex panel as controls to calculate the normal limits with a 95% confidence interval (approved by the Ethics Committee of the Nagasaki University Graduate School of Biomedical Sciences, project registration No.: 14051404). All healthy donors had no past or present medical history of inflammatory disease (*n* = 101 (59 women), mean age 58 (standard deviation: 10) years).

For each cytokine/chemokine, patient values were normalized to healthy controls using Z-scores calculated as follows: Z = (patient value − healthy mean) / healthy SD. This normalization allowed for the direct comparison of different cytokines on a common scale, with a Z-score of $|Z| \leq 2$ defined as within the normal range (corresponding to approximately 95% of healthy individuals).

Changes in cytokine/chemokine Z-scores from baseline to week 52 (ΔZ-scores) within each treatment group were assessed using paired *t*-tests. Between-group comparisons of ΔZ-scores were performed using Welch's *t*-test, which does not assume equal variances. Pearson's

correlation coefficients were used to evaluate inter-cytokine/chemokine relationships and the relationships between cytokines/chemokines and clinical parameters at baseline across all patients. Given the exploratory nature of the study, statistical significance was set at $p < 0.05$ without multiple testing corrections. All analyses were performed using R (version 4.5.0; R Foundation for Statistical Computing, Vienna, Austria).

3. Results and Discussion

The baseline characteristics of both treatment groups are summarized in Table 1. The filgotinib cohort included five patients (median age 60 years, range 37-61; three males) with plasma-cell-type iMCD. The tocilizumab cohort also included five patients (median age 62 years, range 43-69; three males). While all patients receiving tocilizumab were treatment-naïve, two patients receiving filgotinib had prior prednisolone exposure, and one was receiving concomitant prednisolone (10 mg/day). Both groups had similar baseline disease activity, as reflected by the median CHAP score of 5 in both groups. In the filgotinib group, CRP decreased from 6.73 to 3.30 mg/dL and albumin increased from 2.9 to 3.3 g/dL, while CRP decreased from 4.90 to 0.01 mg/dL, hemoglobin increased from 9.3 to 14.1 g/dL, albumin increased from 2.4 to 4.2 g/dL, and CHAP score improved from 5 to 0 in the tocilizumab group.

Figure 1 shows the time series of 41 cytokines and chemokines and clinical parameters (CRP, hemoglobin, albumin, and CHAP score). At baseline, both treatment groups showed marked elevation of multiple cytokines compared to healthy controls (Figure 1, green shaded area: mean ± 2 SD of healthy control). The most prominently elevated cytokines included IL-12p70 (Z = 31.7), IL-22 (Z = 10.0), IFN-γ (Z = 9.5), IL-6 (Z = 9.1), and TNF-β (Z = 7.3) (Figure 2A). Additional notably elevated cytokines and chemokines included GM-CSF (Z = 4.4), TNF-α (Z = 3.7), MCP-3 (Z = 3.1), Fractalkine (Z = 3.0), FGF-2 (Z = 2.7), PDGF-AA (Z = 2.5), and IL-12p40 (Z = 2.1).

3.1. Cytokine/chemokine correlations at baseline

We analyzed the correlations among 41 serum cytokines at baseline in 10 patients. Pearson's correlation analysis revealed several statistically significant cytokine clusters (Figure 2B). IL-6 showed significant positive correlations with TNF-α and TNF-β levels. IL-13 and IL-22 also had significant positive correlations with IL-6, TNF-α, and TNF-β (lower right corner of Figure 2B). Meanwhile, IL-10, TGF-α, IL-15, and IL-1α had significant positive correlations with each other (upper left corner of Figure 2B). These separate correlation clusters suggest the presence of multiple inflammatory axes within the iMCD cytokine/chemokine network that may respond differently to targeted therapies.

3.2. Cytokine/chemokine correlations with clinical parameters at baseline

We analyzed the correlations between cytokine/chemokine and clinical parameters (Figure 2C). CRP levels at baseline were significantly negatively correlated with G-CSF levels. Meanwhile, baseline hemoglobin and albumin levels were significantly positively correlated with MCP-1 and IL-4 levels, respectively. No cytokines or chemokines were significantly correlated with the CHAP score at baseline.

3.3. Cytokine/chemokine changes in two groups

In the filgotinib group, paired *t*-tests identified 12 cytokines with statistically significant changes from baseline to week 52 ($p < 0.05$, Table 2, Figure 2D). The most significant decreases were observed for IL-6 (change of Z score from baseline to 52 weeks (ΔZ)= − 2.24, $p = 0.001$), IL-1RA (ΔZ = − 0.44, $p = 0.003$), and IFN-γ (ΔZ = − 5.12, $p = 0.005$). Additional cytokines that showed significant decreases included IL-2, FGF-2, MCP-3, IL-15, VEGF-A, EGF, IL-17A, and IL-12p40. MCP-1 showed a significant increase (ΔZ = + 0.45, $p = 0.013$). In the tocilizumab group, paired *t*-tests identified 12 cytokines with statistically significant changes ($p < 0.05$, Table 2, Figure 2D). PDGF-AA showed the most significant decrease (ΔZ = − 1.89, $p = 0.011$), followed by MCP-3 (ΔZ = − 3.50, $p = 0.013$), VEGF-A (ΔZ = − 2.11, $p = 0.023$). Additional cytokines that showed significant decreases included IL-4, IL-17F, TGF-α, IL-22, TNF-β, FLT-3L, and FGF-2. As expected with IL-6 receptor blockade, serum IL-6 levels showed a significant increase (ΔZ = + 126.77, $p = 0.047$) due to impaired receptor-mediated clearance. MCP-1 also increased significantly (ΔZ = +1.00, $p = 0.029$).

3.4. Between-group comparison of cytokine/chemokine changes

A direct comparison of ΔZ-scores between treatment groups using Welch's *t*-test identified only three of 41 cytokines with statistically significant differences ($p < 0.05$, Figure 2E). IL-6 showed the largest difference (ΔZ: − 2.24 in filgotinib *vs.* + 126.8 in tocilizumab, $p = 0.045$), reflecting the pharmacological artifact of IL-6 accumulation following receptor blockade. PDGF-AA showed greater suppression with tocilizumab (ΔZ: − 0.50 *vs.* −1.89, $p = 0.028$), whereas IL-15 showed greater suppression with filgotinib (ΔZ: − 0.94 *vs.* −0.12, $p = 0.040$).

3.5. Overall cytokine/chemokine change comparison

When comparing the overall cytokine/chemokine changes between the treatments, both achieved equivalent effects on cytokine/chemokine changes (Figure 2F).

Figure 1. Time-course changes in serum cytokines/chemokines and clinical parameters in patients with iMCD treated with filgotinib or tocilizumab. Serum levels of 41 cytokines/chemokines are shown for patients treated with filgotinib (blue line) and tocilizumab (red dashed line) during the treatment period. The green shaded area indicates the normal range (mean (dashed green lines) ± 2 SD of healthy controls, n = 101). The lower panels show the corresponding clinical parameters: C-reactive protein (CRP), hemoglobin (Hb), albumin (Alb), and CHAP score. For the filgotinib cohort, measurements were obtained at baseline and at weeks 4, 8, 18, 26, 34, and 52. For the tocilizumab cohort, measurements were obtained at baseline and post-treatment (Post-Tx; median, 28 months; range, 11-73 months).

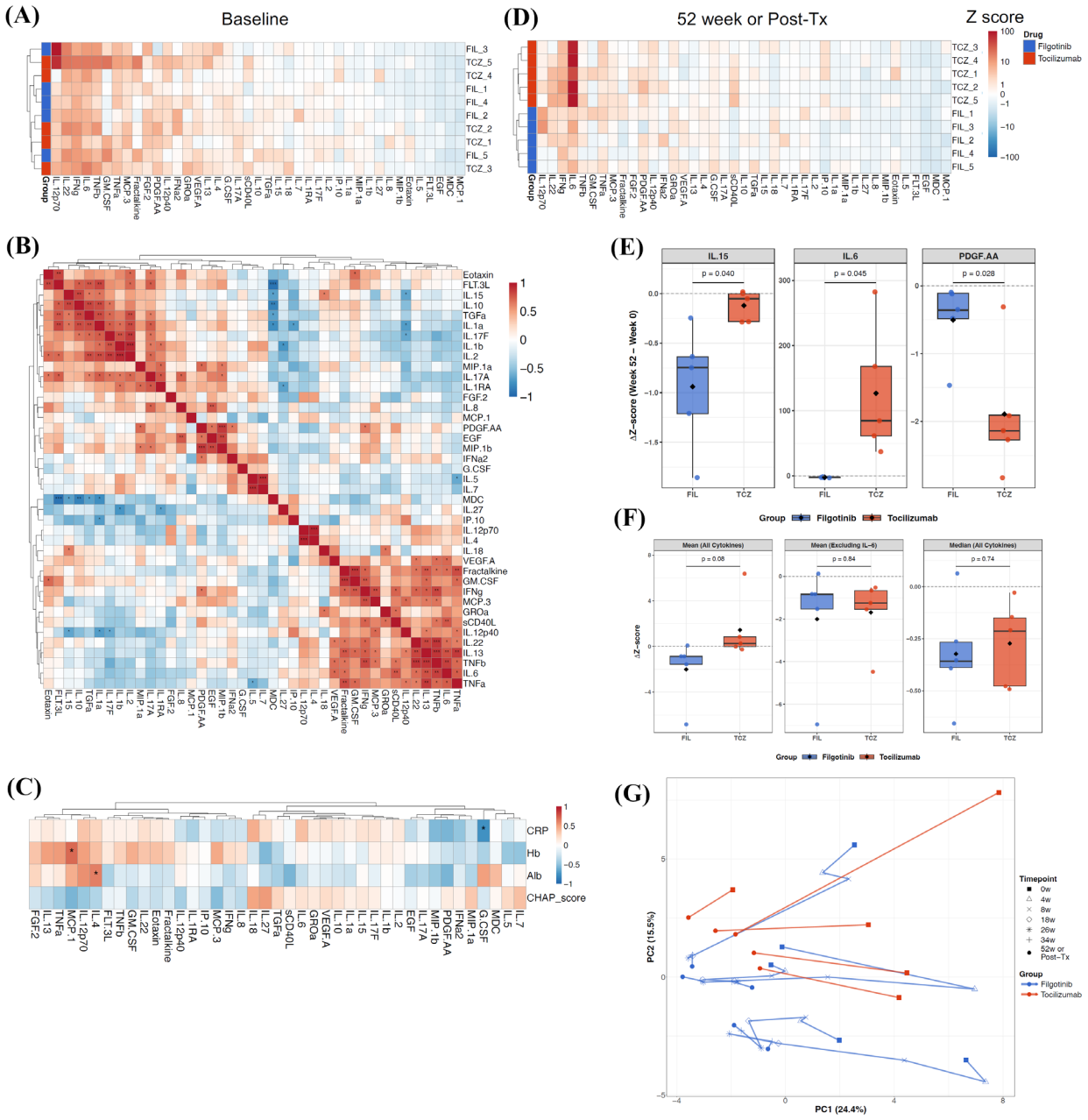

Figure 2. Comprehensive cytokine/chemokine analysis comparing filgotinib and tocilizumab treatment in iMCD. (**A**) Z-score-normalized baseline cytokine/chemokine profiles of all 10 patients. Blue: filgotinib, red: tocilizumab. (**B**) Correlation matrix of the 41 cytokines/chemokines at baseline across all patients. Colors indicate Pearson's correlation coefficients (red: positive correlation; blue: negative correlation). Asterisks denote statistical significance (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). (**C**) Correlation between baseline cytokine/chemokine levels and clinical parameters (CRP, hemoglobin, albumin, and CHAP scores). Significant correlations are indicated by asterisks (*$p < 0.05$). (**D**) Z-score-normalized cytokine/chemokine profiles of all 10 patients at 52 weeks. Blue: filgotinib, red: tocilizumab. (**E**) Between-group comparison of ΔZ-scores for all 41 cytokines/chemokines. IL-6, IL-15, and PDGF-AA levels were significantly different between the treatments (Welch's *t*-test, $p < 0.05$). (**F**) Overall comparison of cytokine/chemokine changes between the treatment groups. Box plots show the distribution of ΔZ-scores across all cytokines for each treatment, with and without IL-6. (**G**) Principal component analysis (PCA) trajectory plot of the 41 cytokines/chemokines. Arrows indicate the treatment trajectories from baseline to the endpoint for each patient. Filgotinib-treated patients are shown with solid blue arrows (7 timepoints); tocilizumab-treated patients are shown with dashed red arrows (2 timepoints).

Using all 41 cytokines/chemokines with mean ΔZ-scores, filgotinib showed − 2.01 ± 2.76 versus tocilizumab + 1.44 ± 2.79 ($p = 0.085$), but this difference was driven entirely by the IL-6 artifact in the tocilizumab group. When IL-6 was excluded from the analysis, the mean ΔZ-scores were also equivalent: − 2.01 ± 2.83 versus − 1.70 ± 1.61 ($p = 0.84$). Using median ΔZ-scores, which are robust to outliers, both groups showed comparative cytokine/chemokine changes(− 0.32 ± 0.26 (filgotinib) *vs.* −0.27 ± 0.21 (tocilizumab), $p = 0.74$). These findings suggest that JAK1 inhibition and IL-6 receptor blockade may achieve comparable broad-spectrum cytokine suppression in iMCD, despite differences in their molecular mechanisms of action.

Table 2. Statistically significant within-group cytokine changes (paired *t*-test, $p < 0.05$)

Filgotinib					Tocilizumab				
Cytokine/chemokine	Baseline	At 52 weeks	ΔZ	*p* value	Cytokine/chemokine	Baseline	Post-treatment	ΔZ	*p* value
IL-6	6.64	4.40	-2.24	0.001	PDGF-AA	3.88	1.99	-1.89	0.011
IL-1RA	0.28	-0.16	-0.44	0.003	MCP-3	4.37	0.87	-3.50	0.013
IFN-γ	7.49	2.38	-5.12	0.005	VEGF-A	1.60	-0.51	-2.11	0.023
MCP-1	-1.32	-0.86	0.45	0.013	IL-4	0.68	0.07	-0.61	0.025
IL-2	0.29	-0.16	-0.45	0.021	MCP-1	-1.43	-0.43	1.00	0.029
FGF-2	3.25	0.01	-3.24	0.026	IL-17F	0.04	-0.10	-0.14	0.030
MCP-3	1.82	0.20	-1.62	0.027	TGF-α	0.34	-0.05	-0.39	0.033
IL-15	0.69	-0.25	-0.94	0.027	IL-22	11.30	3.52	-7.78	0.037
VEGF-A	1.32	0.65	-0.67	0.030	TNF-β	9.79	1.26	-8.54	0.041
EGF	-1.40	-1.50	-0.10	0.032	FLT-3L	-0.67	-0.76	-0.09	0.043
IL-17A	0.65	-0.02	-0.67	0.038	FGF-2	2.13	0.52	-1.62	0.046
IL-12p40	1.35	0.16	-1.19	0.043	IL-6	11.49	138.27	126.77	0.047

ΔZ = Week 52 (filgotinib) or Post-treatment (tocilizumab) Z-score − Baseline Z-score.

3.6. Sensitivity analysis

To address the potential confounding effect of concomitant prednisolone use in one filgotinib-treated patient, we performed a sensitivity analysis excluding this patient (filgotinib $n = 4$ *vs*. tocilizumab $n = 5$). Between-group comparison identified three cytokines with statistically significant differences: FLT-3L (ΔZ: + 0.11 in filgotinib *vs.* − 0.09 in tocilizumab, $p = 0.002$), PDGF-AA (ΔZ: − 0.54 *vs.* − 1.89, $p = 0.039$), and IL-6 (ΔZ: − 2.08 *vs.* + 126.8, $p = 0.045$). IL-15 (ΔZ: − 0.71 *vs.* − 0.12, p = 0.053) became borderline. The overall cytokine suppression equivalence was maintained: mean ΔZ-score − 2.12 ± 3.18 (filgotinib) *vs.* + 1.44 ± 2.79 (tocilizumab) ($p = 0.13$); mean ΔZ-score excluding IL-6: − 2.12 ± 3.25 *vs.* − 1.70 ± 1.61 ($p = 0.82$); median ΔZ-score − 0.31 ± 0.30 *vs.* − 0.27 ± 0.21 ($p = 0.82$). Within-group significant cytokines in the filgotinib group decreased from 12 to 8, consistent with reduced statistical power. The detailed within-group comparison is presented in Supplementary Table S1 (*https://www.ddtjournal.com/action/getSupplementalData.php?ID=293*).

3.7. Principal component analysis with trajectories

Principal component analysis (PCA) of the 41 cytokines and chemokines revealed treatment trajectories in the multidimensional cytokine space (Figure 2G). PC1 and PC2 accounted for 24.4% and 15.5% of the total variance, respectively. At baseline, patients in both treatment groups clustered in the upper-right quadrant, reflecting an elevated inflammatory cytokine state. Over the course of treatment, both filgotinib-treated (seven time points, solid blue arrows) and tocilizumab-treated patients (two time points, dashed red arrows) demonstrated directional movement toward the center-left region of the plot, representing the normalization of cytokine profiles toward healthy control values. The treatment trajectories were largely parallel between the two groups, indicating that JAK1 inhibition and IL-6 receptor blockade induce similar patterns of cytokine modulation, despite their distinct molecular mechanisms of action. PCA trajectory analysis provides compelling visual evidence that filgotinib and tocilizumab achieve comparable immunomodulatory effects in iMCD from the viewpoint of serum cytokines and chemokines.

A notable and clinically important finding was the differential clinical response between the two treatment groups, despite equivalent cytokine/chemokine suppression. This disparity likely reflects the distinct mechanisms of action of these therapies and the specific role of IL-6 in the regulation of acute-phase proteins.

IL-6 is the primary driver of hepatic acute-phase protein synthesis, including CRP, serum amyloid A, and fibrinogen (*19*). IL-6 induces the production of positive acute-phase proteins while simultaneously inhibiting the synthesis of negative acute-phase proteins, such as albumin and transferrin (*20*). IL-6 also suppresses erythropoiesis by promoting hepcidin synthesis, resulting in hypoferremia and anemia (*21*). Tocilizumab achieves complete IL-6 receptor blockade by competitively inhibiting IL-6 binding to both membrane-bound and soluble IL-6 receptors (*22*), thereby comprehensively suppressing the acute phase response. In contrast, filgotinib inhibits JAK1-mediated signaling downstream of the IL-6 receptor but does not completely block the pathway, which may lead to residual IL-6 signaling sufficient to maintain elevated CRP levels, suppress albumin synthesis, and cause anemia. Additionally, the shorter duration of follow-up in the filgotinib cohort (52 weeks) compared with that in the tocilizumab cohort (median 28 months) may have contributed to the observed differences.

A direct comparison of ΔZ-scores between treatment groups using Welch's *t*-test identified only three of the 41 cytokines, IL-6, PDGF-AA, and IL-15, with statistically significant differences (Figure 2E). Because the elevation of IL-6 by tocilizumab is a well-known artifact, we concentrated on IL-15 and PDGF-AA.

The preferential suppression of IL-15 by filgotinib

is consistent with its mechanism of action as a JAK1 inhibitor and provides important insights into the broader immunomodulatory effects of JAK1 inhibition. Upon IL-15 binding, JAK1 and JAK3 are recruited, leading to the phosphorylation of STAT3 *via* JAK1 and STAT5 *via* JAK3 (*23*). In addition, the correlation between IL-15 and other pro-inflammatory cytokines (IL-10, TGF-α, IL-1α) observed in our baseline analysis suggests that IL-15 may participate in a distinct inflammatory axis within the iMCD cytokine network that operates independently of the IL-6 pathway. However, the clinical significance of IL-15 suppression by filgotinib remains to be elucidated. Despite achieving substantial IL-15 reduction, patients treated with filgotinib showed only modest improvements in CRP, hemoglobin, and albumin levels compared to the apparent normalization observed with tocilizumab. This dissociation between IL-15 suppression and modest clinical response suggests that IL-15 may not be a critical driver of core pathophysiology in iMCD with plasma-cell-type histology.

Conversely, the greater suppression of PDGF-AA by tocilizumab warrants careful consideration in the context of iMCD pathophysiology. PDGF-AA transduces potent mitogenic signals and induces actin reorganization (*24*). In iMCD, PDGF-AA may contribute to the characteristic lymph node histopathology and vascular proliferation observed in the affected tissues. However, because serum PDGF-AA levels in iMCD were comparable to those in healthy controls in a previous study (*25*), further research on the role of PDGF-AA in the pathogenesis of iMCD is warranted.

This study has several important limitations that must be considered when interpreting the results. First, the small sample size (n = 5 per group) severely limits statistical power and the generalizability of the findings. Second, the non-randomized design—combining a prospective single-arm trial (filgotinib) with a retrospective cohort (tocilizumab)—introduces potential selection bias. Third, the mismatched follow-up durations between the two cohorts (52 weeks *vs.* a median of 28 months) may have contributed to the more pronounced clinical improvement observed in the tocilizumab group as longer treatment exposure could allow for greater therapeutic benefit. Fourth, concomitant prednisolone use in one filgotinib-treated patient (10 mg/day) represents a confounding factor, as prednisolone possesses broad anti-inflammatory and cytokine-modulating properties that could obscure the true treatment effect of filgotinib. Although sensitivity analyses excluding this patient confirmed the robustness of the primary findings, the contribution of concomitant prednisolone use cannot be entirely excluded in the primary analysis. Given these constraints, the present findings should be considered exploratory and hypothesis-generating. Large-scale, prospective, randomized controlled trials are warranted to validate these observations.

In conclusion, this exploratory analysis from a small-sample study suggests that filgotinib may achieve broad-spectrum cytokine and chemokine suppression comparable to that of tocilizumab in patients with iMCD. The discordance between cytokine/chemokine improvement and clinical parameter improvement suggests the need to identify the important drivers of iMCD pathogenesis in addition to IL-6.

Acknowledgements

The authors thank the patients who participated in this study and their families for their cooperation.

Funding: This study was sponsored by Gilead Sciences, Inc. (CO-JP-986-7009), which provided funding and filgotinib. The funding source had no role in the study design, data collection, analysis, interpretation, manuscript writing, or decision to submit for publication.

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Fajgenbaum DC, Uldrick TS, Bagg A, *et al*. International, evidence-based consensus diagnostic criteria for HHV-8-negative/idiopathic multicentric Castleman disease. Blood. 2017; 129:1646-1657.
2. Fajgenbaum DC, van Rhee F, Nabel CS. HHV-8-negative, idiopathic multicentric Castleman disease: novel insights into biology, pathogenesis, and therapy. Blood. 2014; 123:2924-2933.
3. Mukherjee S, Martin R, Sande B, Paige JS, Fajgenbaum DC. Epidemiology and treatment patterns of idiopathic multicentric Castleman disease in the era of IL-6-directed therapy. Blood Adv. 2022; 6:359-367.
4. Fujimoto S, Koga T, Kawakami A, *et al*. Tentative diagnostic criteria and disease severity classification for Castleman disease: A report of the research group on Castleman disease in Japan. Mod Rheumatol. 2018; 28:161-167.
5. Yoshizaki K, Matsuda T, Nishimoto N, Kuritani T, Taeho L, Aozasa K, Nakahata T, Kawai H, Tagoh H, Komori T, Kishimoto S, Hirano T, Kishimoto T. Pathogenic significance of interleukin-6 (IL-6/BSF-2) in Castleman's disease. Blood. 1989; 74:1360-1367.
6. Chan JY, Loh JW, Lim JQ, *et al.* Single-cell landscape of idiopathic multicentric Castleman disease in identical twins. Blood. 2024; 143:1837-1844.
7. Nishimoto N, Kanakura Y, Aozasa K, *et al.* Humanized anti-interleukin-6 receptor antibody treatment of multicentric Castleman disease. Blood. 2005; 106:2627-2632.
8. Van Rhee F, Wong RS, Munshi N, *et al.* Siltuximab for multicentric Castleman's disease: a randomised, double-blind, placebo-controlled trial. Lancet Oncol. 2014; 15:966-974.
9. van Rhee F, Voorhees P, Dispenzieri A, *et al*. International, evidence-based consensus treatment guidelines for idiopathic multicentric Castleman disease. Blood. 2018; 132:2115-2124.

10. Pierson SK, Stonestrom AJ, Shilling D, Ruth J, Nabel CS, Singh A, Ren Y, Stone K, Li H, van Rhee F, Fajgenbaum DC. Plasma proteomics identifies a "chemokine storm" in idiopathic multicentric Castleman disease. Am J Hematol. 2018; 93:902-912.
11. O'Shea JJ, Schwartz DM, Villarino AV, Gadina M, McInnes IB, Laurence A. The JAK-STAT pathway: Impact on human disease and therapeutic intervention. Annu Rev Med. 2015; 66:311-328.
12. Pierson SK, Shenoy S, Oromendia AB, *et al*. Discovery and validation of a novel subgroup and therapeutic target in idiopathic multicentric Castleman disease. Blood Adv. 2021; 5:3445-3456.
13. Fukui S, Sumiyoshi R, Koga T, *et al*. A Phase Ib Investigator-Initiated Trial of Filgotinib in Patients With Idiopathic Multicentric Castleman Disease. Cureus. 2025; 17:e78865.
14. Tanaka Y, Kavanaugh A, Wicklund J, McInnes IB. Filgotinib, a novel JAK1-preferential inhibitor for the treatment of rheumatoid arthritis: An overview from clinical trials. Mod Rheumatol. 2022; 32:1-11.
15. Genovese MC, Kalunian K, Gottenberg JE, Mozaffarian N, Bartok B, Matzkies F, Gao J, Guo Y, Tasset C, Sundy JS, de Vlam K, Walker D, Takeuchi T. Effect of Filgotinib *vs* Placebo on Clinical Response in Patients With Moderate to Severe Rheumatoid Arthritis Refractory to Disease-Modifying Antirheumatic Drug Therapy: The FINCH 2 Randomized Clinical Trial. JAMA. 2019; 322:315-325.
16. Feagan BG, Danese S, Loftus EV Jr, *et al.* Filgotinib as induction and maintenance therapy for ulcerative colitis (SELECTION): a phase 2b/3 double-blind, randomised, placebo-controlled trial. Lancet. 2021; 397:2372-2384.
17. Fukui S, Sumiyoshi R, Koga T, Hosogaya N, Narita S, Morimoto S, Yano H, Katsube A, Yano S, Masaki Y, Tsunoda S, Sato S, Migita K, Kawakami A. Dynamics of Serum Cytokines and Chemokines in Patients With Idiopathic Multicentric Castleman Disease: From a Phase Ib Investigator-Initiated Trial of Filgotinib. Cureus. 2025; 17:e78974.
18. Nishimura Y, Fajgenbaum DC, Pierson SK, *et al*. Validated international definition of the thrombocytopenia, anasarca, fever, reticulin fibrosis, renal insufficiency, and organomegaly clinical subtype (TAFRO) of idiopathic multicentric Castleman disease. Am J Hematol. 2021; 96:1241-1252.
19. Castell JV, Gomez-Lechon MJ, David M, Andus T, Geiger T, Trullenque R, Fabra R, Heinrich PC. Interleukin-6 is the major regulator of acute phase protein synthesis in adult human hepatocytes. FEBS Lett. 1989; 242:237-239.
20. Heinrich PC, Castell JV, Andus T. Interleukin-6 and the acute phase response. Biochem J. 1990; 265:621-636.
21. Nemeth E, Rivera S, Gabayan V, Keller C, Taudorf S, Pedersen BK, Ganz T. IL-6 mediates hypoferremia of inflammation by inducing the synthesis of the iron regulatory hormone hepcidin. J Clin Invest. 2004; 113:1271-1276.
22. Mihara M, Kasutani K, Okazaki M, Nakamura A, Kawai S, Sugimoto M, Matsumoto Y, Ohsugi Y. Tocilizumab inhibits signal transduction mediated by both mIL-6R and sIL-6R, but not by the receptors of other members of IL-6 cytokine family. Int Immunopharmacol. 2005; 5:1731-1740.
23. Fehniger TA, Caligiuri MA. Interleukin 15: biology and relevance to human disease. Blood. 2001; 97:14-32.
24. Heldin CH, Westermark B. Mechanism of action and *in vivo* role of platelet-derived growth factor. Physiol Rev. 1999; 79:1283-1316.
25. Iwaki N, Gion Y, Kondo E, Kawano M, Masunari T, Moro H, Nikkuni K, Takai K, Hagihara M, Hashimoto Y, Yokota K, Okamoto M, Nakao S, Yoshino T, Sato Y. Elevated serum interferon gamma-induced protein 10 kDa is associated with TAFRO syndrome. Sci Rep. 2017; 7:42316.

Received January 3, 2026; Revised March 23, 2026; Accepted March 29, 2026.

**Address correspondence to:*
Atsushi Kawakami, Department of Immunology and Rheumatology, Nagasaki University Graduate School of Biomedical Sciences, 1-7-1 Sakamoto, Nagasaki 852-8501, Japan.
E-mail: atsushik@nagasaki-u.ac.jp

Released online in J-STAGE as advance publication March 31, 2026.

Brief Report

DOI: 10.5582/ddt.2025.01135

Trends in the prescription of constipation medications in Japan (fiscal years 2019-2023): A nationwide baseline study prior to the 2023 clinical guidelines

Hiroyuki Tanaka*, Toshihiro Ishii

Department of Practical Pharmacy, Faculty of Pharmaceutical Sciences, Toho University, Chiba, Japan.

SUMMARY: Constipation is a common gastrointestinal disorder that markedly affects health and quality of life. In Japan, magnesium-based and contact laxatives have long been widely prescribed. Since the launch of lubiprostone in November 2012, several new agents with novel mechanisms have been introduced, and their use has increased. However, to our knowledge, no nationwide study has examined their use. This study aimed to analyze nationwide trends and regional differences in the prescription of medications for constipation. We used data from the Japanese National Health Insurance Claims and Specific Health Checkup Database Open Data from fiscal years (FYs) 2019 to 2023, covering the period when all currently available constipation medications were on the market in Japan, and prior to the inclusion of the drug selection flowchart in the Evidence-Based Clinical Guidelines for Chronic Constipation 2023. Prescription trends were evaluated using defined daily doses per 1,000 inhabitants per day (DID), and regional variations were assessed using standardized claim ratios (SCRs). The DID for contact laxatives decreased during the study period, whereas the DID for other constipation medications increased. Magnesium oxide showed the smallest regional variation, whereas polyethylene glycol and linaclotide showed the largest variations in FY 2019 and FY 2023, respectively. Throughout the 5-year period, the highest SCRs for contact laxatives were observed in Akita, Iwate, and Aomori prefectures. Our findings provide a comprehensive nationwide picture of evolving constipation treatment patterns and regional variations in Japan, establishing baseline data for evaluating the impact of guideline dissemination and evidence accumulation on clinical practice.

Keywords: constipation, contact laxatives, daily dose per 1,000 inhabitants per day, standardized claim ratio

1. Introduction

Constipation is a condition in which defecation is difficult, unsatisfactory, or infrequent. While constipation is a commonly occurring condition, chronic constipation may affect long-term prognosis, as it is associated with an increased risk of cardiovascular disease and mortality (*1-3*) as well as an elevated risk of developing Parkinson's disease (*4*) and kidney disease (*5*). In Japan, magnesium-based and contact laxatives have been predominantly used for the treatment of constipation. However, since the launch of lubiprostone in November 2012, multiple therapeutic agents with novel mechanisms of action have been developed, and their use has increased.

In the Evidence-Based Clinical Guidelines for Chronic Constipation 2023, published by the Japanese Society of Gastroenterology in July 2023, a flowchart of the diagnosis of chronic constipation and the associated therapeutic drug selection was presented (*6*). Prior to this publication, physicians provided empirical treatments in the absence of clearly defined treatment algorithms. In addition, previous surveys investigating the actual usage of constipation medications in Japan have suggested regional variations in the use of magnesium-based and contact laxatives (*7*). However, to date, no study has been conducted on the actual usage patterns of lubiprostone or subsequently launched medications, including linaclotide, elobixibat, naldemedine, and polyethylene glycol.

Recently, the Ministry of Health, Labour and Welfare (MHLW) published the Japanese National Health Insurance Claims and Specific Health Checkup Database Open Data (NDB-OD) online, providing several summary files based on the NDB data. The NDB-OD (including prescription data from April 1, 2014, to March 31, 2015) was first released in October 2016 and has been updated annually thereafter (*8*). Since NDB-OD can be used to understand Japanese healthcare in a simple and comprehensive manner, the number of studies utilizing NDB-OD has been increasing (*6,9,10*).

Furthermore, epidemiological studies using NDB-OD may be useful for monitoring the actual usage patterns of medications (*10*).

This study aimed to examine the real-world use of constipation medications in Japan over a 5-year period from fiscal years (FYs) 2019 to 2023. This study period encompasses both the time before and after the introduction of the drug selection flowchart in the Evidence-Based Clinical Guidelines for Chronic Constipation 2023, which were published in July 2023, and predominantly precedes their publication. This would allow us to establish baseline data for evaluating the impact of guideline dissemination and evidence accumulation on clinical practice.

2. Materials and Methods

2.1. Data source

We obtained the 6th, 7th, 8th, 9th, and 10th NDB-OD datasets from the MHLW website (*https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000177182.html*; accessed July 7, 2025); the collection periods corresponded to FYs 2019 (April 1, 2019, to March 31, 2020), 2020 (April 1, 2020, to March 31, 2021), 2021 (April 1, 2021, to March 31, 2022), 2022 (April 1, 2022, to March 31, 2023), and 2023 (April 1, 2023, to March 31, 2024), respectively. The NDB-OD provides information about pharmaceuticals with high prescription volumes, categorized by three-digit therapeutic classification codes, and low-frequency products are excluded. Furthermore, the actual prescription volumes of oral medications with prescription volumes ≤ 1,000 are anonymized. Two types of files containing data on prescribed oral medications were available: one stratified by sex and 5-year age groups and the other categorized by the 47 prefectures in Japan. In this study, we extracted records for sennoside, sodium picosulfate, magnesium oxide, polyethylene glycol, lubiprostone, linaclotide, elobixibat, and naldemedine prescribed to inpatients and outpatients both within hospitals and at external pharmacies.

2.2. Assessing the time trend of constipation medication use nationwide in Japan

To compare the use of constipation medications in each FY, the defined daily dose (DDD) per 1,000 inhabitants per day (DID) was calculated according to the DDD of the World Health Organization (WHO) (ATC / DDD Index 2025: *https://atcddd.fhi.no/atc_ddd_index/*).

DID = (constipation medications (g, mg, μg) × 1000)/(WHO DDD or assumed DDD (g, mg, μg) × population × 365)

However, the WHO has not established a DDD for sennosides, lubiprostone, or elobixibat, and that for magnesium oxide (7 g) is considerably higher than the standard daily dose in Japan (2 g). Therefore, in this study, the assumed DDDs for these drugs were based on the dosages specified in the Japanese package inserts. The WHO DDDs and assumed DDDs for each constipation medication are presented in Table 1. Population data for each year were obtained from e-Stat (*https://www.e-stat.go.jp/*).

2.3. Constipation medication prescribing across regions

Based on previous reports (*9,11*), the standardized claim ratios (SCRs), adjusted for both sex ratio and age distribution, were calculated as an index showing the level of medical claims in the target prefectures relative to all of Japan (100 for Japan). This index was calculated using the following formula,

SCR = (Observed number of claims / Expected number of claims) × 100,

where the observed number of claims is the total number of claims in a target prefecture in a year, and the expected number of claims is the number of claims in a year if the situation of submitting claims (sex- and age-specific claims rates) for all of Japan is assumed for the population in the target prefecture. The expected number of claims was calculated using the following formula,

Expected number of claims = ∑ [(population of a sex and age group in a target prefecture) × (sex and age specific claim rate of the sex and age group in whole Japan)],

where sex- and age-specific groups comprised 36 groups of various combinations of sex (males, females)

Table 1. Target constipation medications, ATC codes, and defined daily doses

Drugs	WHO ATC code (Classification)	WHO DDD	Daily dosage indicated in the Japanese package insert	WHO DDD or assumed DDD value used to calculate DID
Sennosides	A06AB06 (Contact laxatives)	NA	12 –24 mg	12 mg
Sodium picosulfate	A06AB08 (Contact laxatives)	5 mg	5 –7.5 mg	5 mg
Magnesium oxide	A06AD02 (Osmotically acting laxatives)	7 g	2 g	2 g
Polyethylene glycol	A06AD15 (Osmotically acting laxatives)	10 g	13.125 g	10 g
Lubiprostone	A06AX03 (Other drugs for constipation)	NA	48 μg	48 μg
Linaclotide	A06AX04 (Other drugs for constipation)	0.29 mg	0.5 mg	0.29 mg
Elobixibat	A06AX09 (Other drugs for constipation)	NA	10 mg	10 mg
Naldemedine	A06AH05 (Other drugs for constipation)	0.2 mg	0.2 mg	0.2 mg

WHO, World health Organization; ATC, Anatomical Therapeutic Chemical; DDD, defined daily dose.

and age groups (aged 0-4, 5-9,…., 85 and over).

2.4. Statistical analyses

Trends in the consumption of constipation medications were evaluated using linear regression analysis. Statistical significance was set at $p < 0.05$. To assess regional variation, the coefficient of variation (CV) of SCRs and the Pearson correlation coefficient between FYs 2019 and 2023 were calculated. All statistical analyses were performed using Microsoft Excel 2018 (Microsoft Corp., Redmond, WA, USA) and JMP Student Edition 18 (SAS Institute, Inc., Cary, NC, USA).

2.5. Ethics

The data used in this study were obtained from the NDB-OD, a publicly available and fully anonymized dataset. We consulted our institutional ethics committee regarding the ethical handling of this study. Based on the consultation, it was confirmed that the NDB-OD consists solely of statistical information from which all correspondence to specific individuals has been completely eliminated, and therefore falls outside the scope of the ethical guidelines stipulated by our institution. Accordingly, formal ethical review and informed consent were not required.

3. Results and Discussion

Chronic constipation is a highly prevalent gastrointestinal disorder frequently encountered in gastroenterology and various medical specialties (*12*). Studies investigating the impact of various functional gastrointestinal disorders (including constipation) on survival have reported that chronic constipation is associated with a significantly higher risk of reduced survival than other disorders, warranting careful clinical consideration (*13*). Therefore, standardizing constipation treatment is important, and it is necessary to understand the current situation. The present study elucidated the trends and regional variations in the use of constipation medication in Japan from FY 2019 to FY 2023 using the NDB-OD.

The DIDs for constipation medications are shown in Table 2. Among them, sennosides had the highest DID, followed by magnesium oxide, and this ranking remained consistent over the 5-year period. The use of contact laxatives (sennosides [from 23.58 to 19.39, $p < 0.001$] and sodium picosulfate [from 7.18 to 6.66, $p = 0.016$]) consistently declined, whereas the use of other medications for constipation consistently increased. Polyethylene glycol showed the highest rate of increase in DID (slope = 0.562, $p < 0.001$) among all constipation medications. The DID of naldemedine, which is indicated for opioid-induced constipation (OIC), also showed a modest increase (slope = 0.047, $p = 0.005$).

These findings suggest that the emergence of new drugs for chronic constipation and their increased utilization may have contributed to the decline in the use of contact laxatives. Contact laxatives take effect within a few hours, but they can cause side effects such as watery diarrhea, abdominal pain, dehydration, and electrolyte imbalance (*14*). The long-term use of contact laxatives can also lead to tolerance and dependence; therefore, they should be prescribed with caution (*15*). In addition, the chronic use of sennosides has been reported to be associated with the development of melanosis coli, colorectal adenoma, and/or carcinoma (*16-18*). Therefore, the use of contact laxatives should be kept to a minimum and, whenever possible, administered on an as-needed basis or for short-term use. Magnesium oxide, an osmotically acting laxative, was also frequently used in Japan during the study period, and its DID increased slightly from 17.00 to 17.65 ($p = 0.032$). Cases of hypermagnesemia associated with magnesium oxide use have been reported in patients with impaired renal function and older patients; therefore, it is necessary to appropriately monitor serum magnesium levels and symptoms of hypermagnesemia (*19,20*). Another osmotically acting laxative, polyethylene glycol, also showed an increase in the DID (from 0.70 to 3.00, $p < 0.001$). Polyethylene glycol has been available for prescription in Japan since 2018, and it can be administered to children aged 2 years or older. Moreover, polyethylene glycol is minimally absorbed and does not carry specific warnings for patients with impaired renal function. Polyethylene glycol is a relatively new drug, and its cost is higher than that of magnesium oxide. Therefore, polyethylene glycol is primarily considered a feasible treatment option for patients for whom magnesium oxide

Table 2. Time trends in the nationwide utilization of constipation medications (defined daily doses per 1,000 inhabitants per day)

Drugs	FY 2019	FY 2020	FY 2021	FY 2022	FY 2023	Slope	95% CI	*p*-value
Sennosides	23.58	22.58	21.77	20.63	19.39	-1.033	-1.179 to -0.887	< 0.001
Sodium picosulfate	7.18	7.04	7.08	6.84	6.66	-0.124	-0.206 to -0.042	0.016
Magnesium oxide	17.00	17.24	17.62	17.64	17.65	0.170	0.027 to 0.313	0.032
Polyethylene glycol	0.70	1.44	2.03	2.46	3.00	0.562	0.461 to 0.663	< 0.001
Lubiprostone	2.03	2.14	2.28	2.37	2.51	0.119	0.106 to 0.132	< 0.001
Linaclotide	1.16	1.36	1.56	1.67	1.80	0.159	0.121 to 0.197	< 0.001
Elobixibat	0.77	1.12	1.45	1.73	1.99	0.305	0.269 to 0.341	< 0.001
Naldemedine	0.17	0.19	0.23	0.28	0.36	0.047	0.027 to 0.067	0.005

therapy is inappropriate or insufficiently effective. In recent years, lubiprostone (chloride channel (ClC-2) activator), linaclotide (guanylate cyclase-C agonist), and elobixibat (ileal bile acid transporter inhibitor) have been approved for insurance coverage, thereby expanding the treatment options for chronic constipation. However, an established strategy for selecting drugs for patients with specific clinical characteristics is currently unavailable, and further investigation is required in this regard (*6*). Osmotically acting and contact laxatives are considered the preferred treatments for OIC. Strong evidence supports the efficacy of naldemedine, a peripheral μ-opioid receptor antagonist, in the treatment of OIC (*21*); however, owing to cost-effectiveness concerns, physicians differ in their opinions regarding its use as a first-line therapy (*6*).

The SCRs for each constipation medication in the 47 prefectures during FYs 2019 and 2023 are displayed in a box plot in Figure 1. A heat map illustrating the SCRs for each constipation medication across the 47 prefectures in FYs 2019 and 2023 is presented in Figure 2. Magnesium oxide exhibited the least regional variation. The largest regional differences were observed for polyethylene glycol in FY 2019 and linaclotide in FY 2023. The highest SCR for linaclotide was observed in Tokushima in both years, at 198.4 in FY 2019 and 270.6 in FY 2023. Akita, Iwate, and Aomori had the highest utilization of contact laxatives (sennosides and sodium picosulfate). The CV of SCRs was similar between FYs 2019 and 2023 (CV = 0.129 and 0.123, respectively), and a strong correlation was observed between prefecture-level SCRs in these years (r = 0.964), indicating that the magnitude and pattern of regional variation remained largely unchanged over time. Box plots illustrating the 5-year changes in the SCRs for all constipation medications across each prefecture are presented in Supplementary Figures S1-8 (*https://www.ddtjournal.com/action/getSupplementalData.php?ID=295*).

This study clarifies regional differences in prescribing constipation medications across 47 prefectures in Japan. Previous studies have shown that the prevalence of constipation increases with age in both men and women, and it is higher in women than in men (*7,22*). As such, the SCR was used as an age- and sex-adjusted index in this study. Among the examined constipation medications, the SCR for magnesium oxide exhibited the smallest regional variation. The SCRs for contact laxatives were particularly high in Akita, Iwate, and Aomori, all of which are located at higher latitudes and have lower average temperatures. Previous studies using the NDB-OD have shown that the use of contact laxatives is associated with lower ambient temperature (*7*). Johanson was the first to report cold temperature as a global environmental risk factor for constipation (*23*). Consistent with these findings, the SCRs for all constipation medications

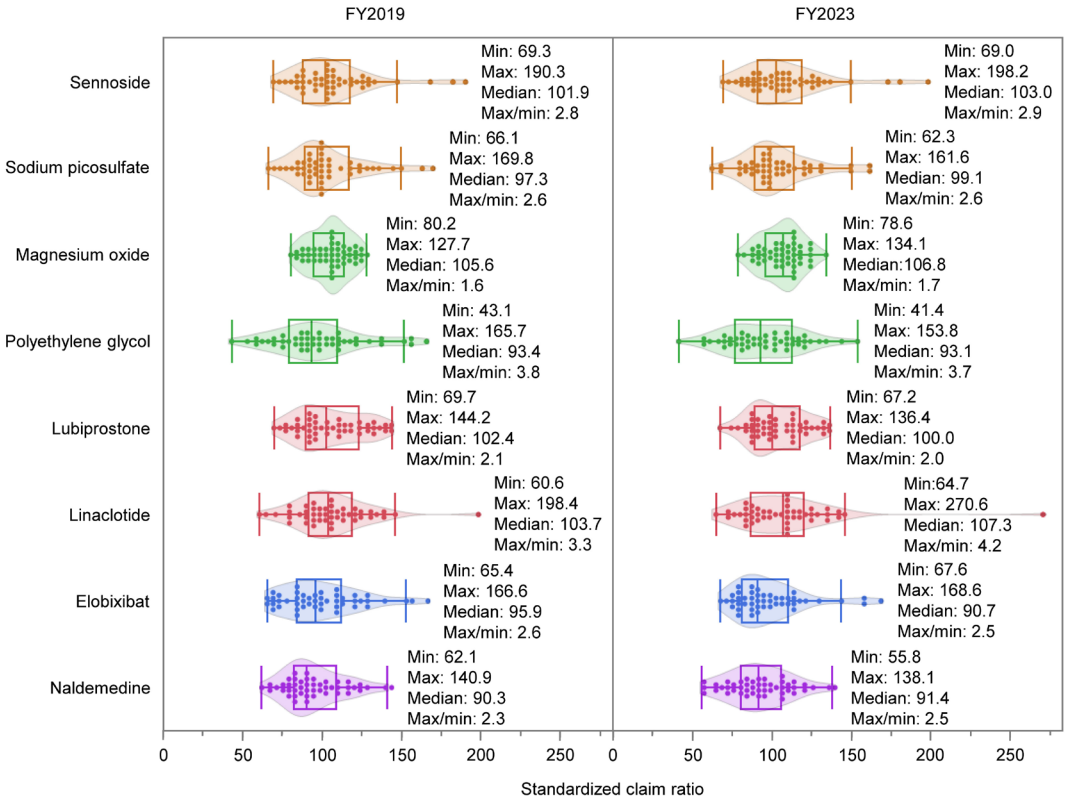

Figure 1. Regional variation in standardized claim ratios for constipation medications across Japan's 47 prefectures: Box plots for FYs 2019 and 2023. Sennosides and sodium picosulfate: contact laxatives; magnesium oxide and polyethylene glycol: osmotically acting laxatives; lubiprostone and linaclotide: intestinal secretagogues; elobixibat: ileal bile acid transporter inhibitor; naldemedine: peripherally acting μ-opioid receptor antagonist.

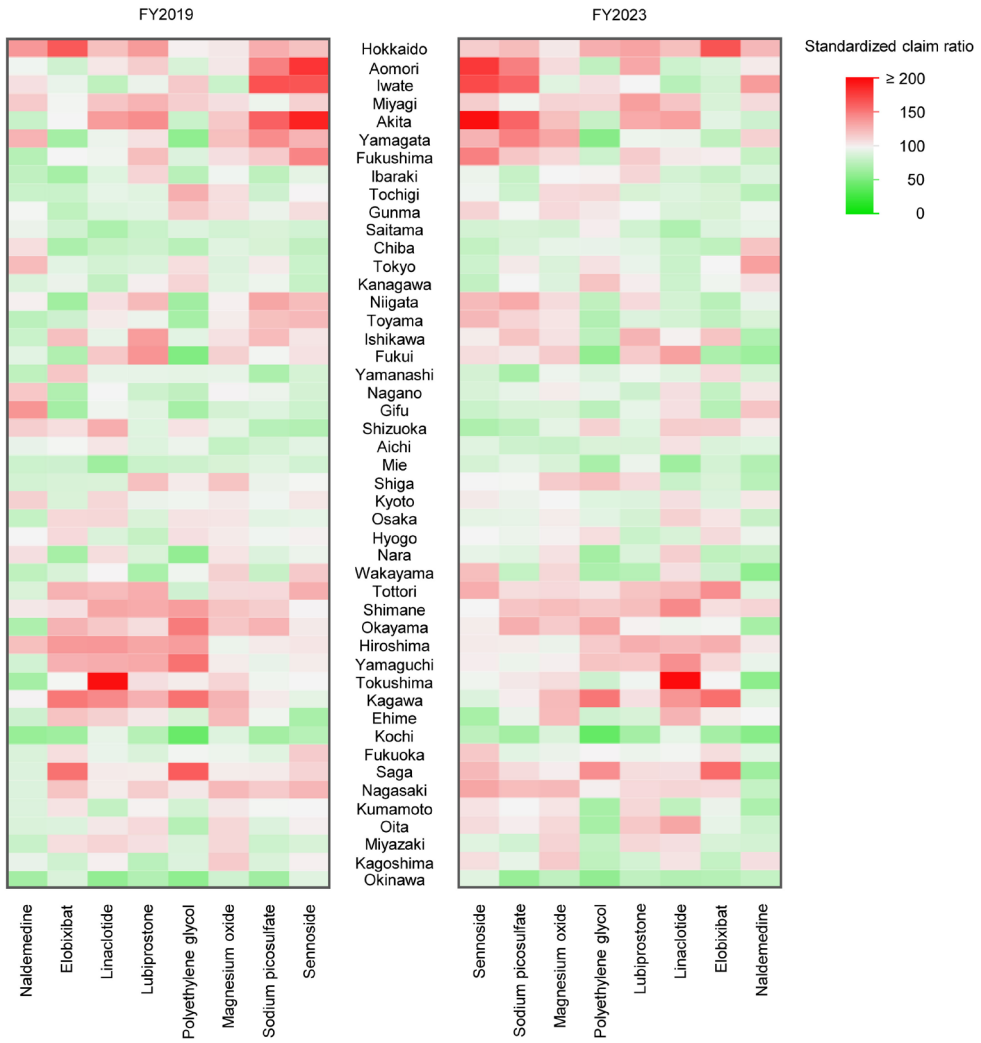

Figure 2. Regional variation in standardized claim ratios for constipation medications across Japan's 47 prefectures: Heat maps for FYs 2019 and 2023.

investigated were > 100 in Hokkaido, which has the lowest average temperature, whereas the SCRs were < 100 for Okinawa, which has the highest average temperature (average temperature data for Japan were obtained from e-Stat [*https://www.e-stat.go.jp/*]).

This study has several limitations. First, the investigation was limited to the eight constipation medications listed in Table 1, and herbal medicines, suppositories, and other constipation-related medications were excluded. Second, for drugs without an established WHO DDD, or those whose WHO DDD substantially differed from the dosages indicated in the Japanese package inserts, assumed DDDs were set specifically for this study. Third, several potential confounding factors, such as other diseases or medications related to constipation and regional characteristics, were not considered in the analysis. Fourth, because this study focused solely on prescribed medications, it did not fully capture the actual burden of constipation in Japan. An internet survey of the general Japanese population reported that among those who had used laxatives, 67.5% had purchased over-the-counter (OTC) drugs from pharmacies (*24*). Therefore, some constipation medications are available both as prescription and OTC drugs, while others are used exclusively as prescription drugs. Consequently, comparing these two groups based on DID was not appropriate.

In conclusion, despite its limitations, this study elucidates the trends and regional differences in the use of constipation medication in Japan from FYs 2019 to 2023. In July 2023, the Evidence-Based Clinical Guidelines for Chronic Constipation 2023 were released in Japan, and these included a drug selection flowchart. The findings of this study are expected to serve as baseline data for future evaluations of the standardization of constipation treatment through the dissemination of guidelines and accumulation of evidence.

Funding: None

Conflict of Interest: The authors have no conflicts of interest to disclose.

References

1. Sumida K, Molnar MZ, Potukuchi PK, Thomas F, Lu JL, Yamagata K, Kalantar-Zadeh K, Kovesdy CP. Constipation and risk of death and cardiovascular events. Atherosclerosis. 2019; 281:114-120.
2. Honkura K, Tomata Y, Sugiyama K, Kaiho Y, Watanabe T, Zhang S, Sugawara Y, Tsuji I. Defecation frequency and cardiovascular disease mortality in Japan: The Ohsaki cohort study. Atherosclerosis. 2016; 246:251-256.
3. Kubota Y, Iso H, Tamakoshi A. Bowel movement frequency, laxative use, and mortality from coronary heart disease and stroke among Japanese men and women: the Japan collaborative cohort (JACC) study. J Epidemiol. 2016; 26:242-248.
4. Adams-Carr KL, Bestwick JP, Shribman S, Lees A, Schrag A, Noyce AJ. Constipation preceding Parkinson's disease: a systematic review and meta-analysis. J Neurol Neurosurg Psychiatry. 2016; 87:710-716.
5. Sumida K, Molnar MZ, Potukuchi PK, Thomas F, Lu JL, Matsushita K, Yamagata K, Kalantar-Zadeh K, Kovesdy CP. Constipation and incident CKD. J Am Soc Nephrol. 2017; 28:1248-1258.
6. Ihara E, Manabe N, Ohkubo H, *et al*. Evidence-Based Clinical Guidelines for Chronic Constipation 2023. Digestion. 2025; 106:62-89.
7. Mihara H, Murayama A, Nanjo S, Ando T, Tajiri K, Fujinami H, Yamada M, Yasuda I. Factors correlated with drug use for constipation: perspectives from the 2016 open Japanese National Database. BMC Gastroenterol. 2020; 20:284.
8. Ministry of Health, Labour and Welfare. The 10th NDB Open Data Commentary. *https://www.mhlw.go.jp/content/12400000/001492909.pdf* (accessed August 9, 2025).
9. Taira K, Mori T, Ishimaru M, Iwagami M, Sakata N, Watanabe T, Takahashi H, Tamiya N. Regional inequality in dental care utilization in Japan: an ecological study using the National Database of Health Insurance Claims. Lancet Reg Health West Pac. 2021; 12:100170.
10. Tanaka H, Onoda T, Ishii T. Understanding the actual use of anti-HIV drugs in Japan from 2016 to 2019: demonstrating epidemiological relevance of NDB Open Data Japan for understanding Japanese medical care. Int J Environ Res Public Health. 2022; 19:12130.
11. Murakami G, Muramatsu K, Hayashida K, Fujimori K, Matsuda S. Ecological analysis of factors associated with mortality of cerebral infarction in Japan. Asian Pac J Dis Manag. 2014; 6:73-76.
12. Camilleri M, Ford AC, Mawe GM, Dinning PG, Rao SS, Chey WD, Simrén M, Lembo A, Young-Fadok TM, Chang L. Chronic constipation. Nat Rev Dis Primers. 2017; 3:17095.
13. Chang JY, Locke GR 3rd, McNally MA, Halder SL, Schleck CD, Zinsmeister AR, Talley NJ. Impact of functional gastrointestinal disorders on survival in the community. Am J Gastroenterol. 2010; 105:822-832.
14. MacLennan WJ, Pooler A. A comparison of sodium picosulphate ("laxoberal") with standardised senna ("senokot") in geriatric patients. Curr Med Res Opin. 1975; 2:641-647.
15. Wald A. Is chronic use of stimulant laxatives harmful to the colon? J Clin Gastroenterol. 2003; 36:386-389.
16. Yamate Y, Hiramoto K, Yokoyama S, Ooi K. Immunological changes in the intestines and skin after senna administration. Pharm Biol. 2015; 53:913-920.
17. Freeman H. "Melanosis" in the small and large intestine. World J Gastroenterol. 2008; 14:4296-4299.
18. Liu ZH, Foo DCC, Law WL, Chan FSY, Fan JKM, Peng JS. Melanosis coli: harmless pigmentation? A case-control retrospective study of 657 cases. PLoS One. 2017; 12:e0186668.
19. Yamaguchi H, Shimada H, Yoshita K, Tsubata Y, Ikarashi K, Morioka T, Saito N, Sakai S, Narita I. Severe hypermagnesemia induced by magnesium oxide ingestion: a case series. CEN Case Rep. 2019; 8:31-37.
20. Ishii H, Sawada R, Shiomi M, Shibuya K. A case-control study showing low creatinine clearance and high magnesium intake as risk factors for hypermagnesemia in older individuals. Magnes Res. 2023; 36:23-30.
21. Esmadi M, Ahmad D, Hewlett A. Efficacy of naldemedine for the treatment of opioid-induced constipation: a meta-analysis. J Gastrointestin Liver Dis. 2019; 28:41-46.
22. Higgins PD, Johanson JF. Epidemiology of constipation in North America: a systematic review. Am J Gastroenterol. 2004; 99:750-759.
23. Johanson JF. Geographic distribution of constipation in the United States. Am J Gastroenterol. 1998; 93:188-191.
24. Kasugai K, Yamamoto S, Kawamura Y, Adachi K, Yamaguchi Y, Tamura Y, Izawa S, Hijikata Y, Ebi M, Funaki Y, Ogasawara N, Sasaki M. Internet survey of the actual situation of constipation in Japanese general population REACTION-J:research for actual situation of constipation in the Japanese. Nihon Shokakibyo Gakkai Zasshi. 2019; 116:913-926. (In Japanese)

Received December 23, 2025; Revised March 18, 2026; Accepted April 11, 2026.

**Address correspondence to:*
Hiroyuki Tanaka, Department of Practical Pharmacy, Faculty of Pharmaceutical Sciences, Toho University, 2-2-1 Miyama, Funabashi, Chiba 274-8510, Japan.
E-mail: hiroyuki.tanaka@phar.toho-u.ac.jp

Released online in J-STAGE as advance publication April 15, 2026.

DD&T
Drug Discoveries & Therapeutics

Guide for Authors

1. Scope of Articles

Drug Discoveries & Therapeutics (Print ISSN 1881-7831, Online ISSN 1881-784X) welcomes contributions in all fields of pharmaceutical and therapeutic research such as medicinal chemistry, pharmacology, pharmaceutical analysis, pharmaceutics, pharmaceutical administration, and experimental and clinical studies of effects, mechanisms, or uses of various treatments. Studies in drug-related fields such as biology, biochemistry, physiology, microbiology, and immunology are also within the scope of this journal.

2. Submission Types

Original Articles should be well-documented, novel, and significant to the field as a whole. An Original Article should be arranged into the following sections: Title page, Abstract, Introduction, Materials and Methods, Results, Discussion, Acknowledgments, and References. Original articles should not exceed 5,000 words in length (excluding references) and should be limited to a maximum of 50 references. Articles may contain a maximum of 10 figures and/or tables. Supplementary Data are permitted but should be limited to information that is not essential to the general understanding of the research presented in the main text, such as unaltered blots and source data as well as other file types.

Brief Reports definitively documenting either experimental results or informative clinical observations will be considered for publication in this category. Brief Reports are not intended for publication of incomplete or preliminary findings. Brief Reports should not exceed 3,000 words in length (excluding references) and should be limited to a maximum of 4 figures and/or tables and 30 references. A Brief Report contains the same sections as an Original Article, but the Results and Discussion sections should be combined.

Reviews should present a full and up-to-date account of recent developments within an area of research. Normally, reviews should not exceed 8,000 words in length (excluding references) and should be limited to a maximum of 10 figures and/or tables and 100 references. Mini reviews are also accepted, which should not exceed 4,000 words in length (excluding references) and should be limited to a maximum of 5 figures and/or tables and 50 references.

Policy Forum articles discuss research and policy issues in areas related to life science such as public health, the medical care system, and social science and may address governmental issues at district, national, and international levels of discourse. Policy Forum articles should not exceed 3,000 words in length (excluding references) and should be limited to a maximum of 5 figures and/or tables and 30 references.

Case Reports should be detailed reports of the symptoms, signs, diagnosis, treatment, and follow-up of an individual patient. Case reports may contain a demographic profile of the patient but usually describe an unusual or novel occurrence. Unreported or unusual side effects or adverse interactions involving medications will also be considered. Case Reports should not exceed 3,000 words in length (excluding references).

Communications are short, timely pieces that spotlight new research findings or policy issues of interest to the field of global health and medical practice that are of immediate importance. Depending on their content, Communications will be published as "Comments" or "Correspondence". Communications should not exceed 1,500 words in length (excluding references) and should be limited to a maximum of 2 figures and/or tables and 20 references.

Editorials are short, invited opinion pieces that discuss an issue of immediate importance to the fields of global health, medical practice, and basic science oriented for clinical application. Editorials should not exceed 1,000 words in length (excluding references) and should be limited to a maximum of 10 references. Editorials may contain one figure or table.

News articles should report the latest events in health sciences and medical research from around the world. News should not exceed 500 words in length.

Letters should present considered opinions in response to articles published in *Drug Discoveries & Therapeutics* in the last 6 months or issues of general interest. Letters should not exceed 800 words in length and may contain a maximum of 10 references. Letters may contain one figure or table.

3. Editorial Policies

For publishing and ethical standards, *Drug Discoveries & Therapeutics* follows the Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals issued by the International Committee of Medical Journal Editors (ICMJE, *https://icmje.org/recommendations*), and the Principles of Transparency and Best Practice in Scholarly Publishing jointly issued by the Committee on Publication Ethics (COPE, *https://publicationethics.org/resources/guidelines-new/principles-transparency-and-best-practice-scholarly-publishing*), the Directory of Open Access Journals (DOAJ, *https://doaj.org/apply/transparency*), the Open Access Scholarly Publishers Association (OASPA, *https://oaspa.org/principles-of-transparency-and-best-practice-in-scholarly-publishing-4*), and the World Association of Medical Editors (WAME, *https://wame.org/principles-of-transparency-and-best-practice-in-scholarly-publishing*).

Drug Discoveries & Therapeutics will perform an especially prompt review to encourage innovative work. All original research will be subjected to a rigorous standard of peer review and will be edited by experienced copy editors to the highest standards.

Ethical Approval of Studies and Informed Consent: For all manuscripts reporting data from studies involving human participants or animals, formal review and approval, or formal review and waiver, by an appropriate institutional review board or ethics committee is required and should be described in the Methods section. When your manuscript contains any case details, personal information and/or images of patients or other individuals, authors must obtain appropriate written consent, permission and release in order to comply with all applicable laws and regulations concerning privacy and/or security of personal information. The consent form needs to comply with the relevant legal requirements of your particular jurisdiction, and please do not send signed consent form to *Drug Discoveries & Therapeutics* to respect your patient's and any other individual's privacy. Please instead describe the information clearly in the Methods (patient consent) section of your manuscript while retaining copies of the signed forms in the event they should be needed. Authors should also state that the study conformed to the provisions of the Declaration of Helsinki (as revised in 2013, *https://wma.net/what-we-do/medical-ethics/declaration-of-helsinki*). When reporting experiments on animals, authors should indicate whether the institutional and national guide for the care and use of laboratory animals was followed.

Reporting Clinical Trials: The ICMJE (*https:// icmje.org/recommendations/browse/publishing-and-editorial-issues/clinical-trial-registration.html*) defines a clinical trial as any research project that prospectively assigns people or a group of people to an intervention, with or without concurrent comparison or control groups, to study the relationship between a health-related intervention and a health outcome. Registration of clinical trials in a public trial registry

at or before the time of first patient enrollment is a condition of consideration for publication in *Drug Discoveries & Therapeutics*, and the trial registration number will be published at the end of the Abstract. The registry must be independent of for-profit interest and publicly accessible. Reports of trials must conform to CONSORT 2010 guidelines (*https://consort-statement.org/consort-2010*). Articles reporting the results of randomized trials must include the CONSORT flow diagram showing the progress of patients throughout the trial.

Conflict of Interest: All authors are required to disclose any actual or potential conflict of interest including financial interests or relationships with other people or organizations that might raise questions of bias in the work reported. If no conflict of interest exists for each author, please state "There is no conflict of interest to disclose".

Submission Declaration: When a manuscript is considered for submission to *Drug Discoveries & Therapeutics*, the authors should confirm that 1) no part of this manuscript is currently under consideration for publication elsewhere; 2) this manuscript does not contain the same information in whole or in part as manuscripts that have been published, accepted, or are under review elsewhere, except in the form of an abstract, a letter to the editor, or part of a published lecture or academic thesis; 3) authorization for publication has been obtained from the authors' employer or institution; and 4) all contributing authors have agreed to submit this manuscript.

Initial Editorial Check: Immediately after submission, the journal's managing editor will perform an initial check of the manuscript. A suitable academic editor will be notified of the submission and invited to check the manuscript and recommend reviewers. Academic editors will check for plagiarism and duplicate publication at this stage. The journal has a formal recusal process in place to help manage potential conflicts of interest of editors. In the event that an editor has a conflict of interest with a submitted manuscript or with the authors, the manuscript, review, and editorial decisions are managed by another designated editor without a conflict of interest related to the manuscript.

Peer Review: *Drug Discoveries & Therapeutics* operates a single-anonymized review process, which means that reviewers know the names of the authors, but the authors do not know who reviewed their manuscript. All articles are evaluated objectively based on academic content. External peer review of research articles is performed by at least two reviewers, and sometimes the opinions of more reviewers are sought. Peer reviewers are selected based on their expertise and ability to provide quality, constructive, and fair reviews. For research manuscripts, the editors may, in addition, seek the opinion of a statistical reviewer. Every reviewer is expected to evaluate the manuscript in a timely, transparent, and ethical manner, following the COPE guidelines (*https://publicationethics.org/files/cope-ethical-guidelines-peer-reviewers-v2_0.pdf*). We ask authors for sufficient revisions (with a second round of peer review, when necessary) before a final decision is made. Consideration for publication is based on the article's originality, novelty, and scientific soundness, and the appropriateness of its analysis.

Suggested Reviewers: A list of up to 3 reviewers who are qualified to assess the scientific merit of the study is welcomed. Reviewer information including names, affiliations, addresses, and e-mail should be provided at the same time the manuscript is submitted online. Please do not suggest reviewers with known conflicts of interest, including participants or anyone with a stake in the proposed research; anyone from the same institution; former students, advisors, or research collaborators (within the last three years); or close personal contacts. Please note that the Editor-in-Chief may accept one or more of the proposed reviewers or may request a review by other qualified persons.

Language Editing: Manuscripts prepared by authors whose native language is not English should have their work proofread by a native English speaker before submission. If not, this might delay the publication of your manuscript in *Drug Discoveries & Therapeutics*.

The Editing Support Organization can provide English proofreading, Japanese-English translation, and Chinese-English translation services to authors who want to publish in *Drug Discoveries & Therapeutics* and need assistance before submitting a manuscript. Authors can visit this organization directly at *https://www.iacmhr.com/iac-eso/support.php?lang=en. IAC-ESO* was established to facilitate manuscript preparation by researchers whose native language is not English and to help edit works intended for international academic journals.

Copyright and Reuse: Before a manuscript is accepted for publication in *Drug Discoveries & Therapeutics*, authors will be asked to sign a transfer of copyright agreement, which recognizes the common interest that both the journal and author(s) have in the protection of copyright. We accept that some authors (*e.g.*, government employees in some countries) are unable to transfer copyright. A JOURNAL PUBLISHING AGREEMENT (JPA) form will be e-mailed to the authors by the Editorial Office and must be returned by the authors by mail, fax, or as a scan. Only forms with a hand-written signature from the corresponding author are accepted. This copyright will ensure the widest possible dissemination of information. Please note that the manuscript will not proceed to the next step in publication until the JPA Form is received. In addition, if excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article.

4. Cover Letter

The manuscript must be accompanied by a cover letter prepared by the corresponding author on behalf of all authors. The letter should indicate the basic findings of the work and their significance. The letter should also include a statement affirming that all authors concur with the submission and that the material submitted for publication has not been published previously or is not under consideration for publication elsewhere. The cover letter should be submitted in PDF format. For an example of Cover Letter, please visit: Download Centre (*https://www.ddtjournal.com/downcentre)*.

5. Submission Checklist

The Submission Checklist should be submitted when submitting a manuscript through the Online Submission System. Please visit Download Centre (*https://www.ddtjournal.com/downcentre)* and download the Submission Checklist file. We recommend that authors use this checklist when preparing your manuscript to check that all the necessary information is included in your article (if applicable), especially with regard to Ethics Statements.

6. Manuscript Preparation

Manuscripts are suggested to be prepared in accordance with the "Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals", as presented at *http://www.ICMJE.org*.

Manuscripts should be written in clear, grammatically correct English and submitted as a Microsoft Word file in a single-column format. Manuscripts must be paginated and typed in 12-point Times New Roman font with 24-point line spacing. Please do not embed figures in the text. Abbreviations should be used as little as possible and should be explained at first mention unless the term is a well-known abbreviation (*e.g.* DNA). Single words should not be abbreviated.

Title page: The title page must include 1) the title of the paper (Please note the title should be short, informative, and contain the major key words); 2) full name(s) and affiliation(s) of the author(s), 3) abbreviated names of the author(s), 4) full name, mailing address, telephone/fax numbers, and e-mail address of the corresponding author; 5) author contribution statements to specify the individual contributions of all authors to this manuscript, and 6) conflicts of interest (if you have an actual or potential conflict of interest to disclose, it must be included as a footnote on the title page of the manuscript; if no conflict of interest

exists for each author, please state "There is no conflict of interest to disclose").

Abstract: The abstract should briefly state the purpose of the study, methods, main findings, and conclusions. For article types including Original Article, Brief Report, Review, Policy Forum, and Case Report, a one-paragraph abstract consisting of no more than 250 words must be included in the manuscript. For Communications, Editorials, News, or Letters, a brief summary of main content in 150 words or fewer should be included in the manuscript. For articles reporting clinical trials, the trial registration number should be stated at the end of the Abstract. Abbreviations must be kept to a minimum and non-standard abbreviations explained in brackets at first mention. References should be avoided in the abstract. Three to six key words or phrases that do not occur in the title should be included in the Abstract page.

Introduction: The introduction should provide sufficient background information to make the article intelligible to readers in other disciplines and sufficient context clarifying the significance of the experimental findings.

Materials/Patients and Methods: The description should be brief but with sufficient detail to enable others to reproduce the experiments. Procedures that have been published previously should not be described in detail but appropriate references should simply be cited. Only new and significant modifications of previously published procedures require complete description. Names of products and manufacturers with their locations (city and state/country) should be given and sources of animals and cell lines should always be indicated. All clinical investigations must have been conducted in accordance with the Declaration of Helsinki (as revised in 2013, *https://wma.net/what-we-do/medical-ethics/declaration-of-helsinki*). All human and animal studies must have been approved by the appropriate institutional review board(s) and a specific declaration of approval must be made within this section.

Results: The description of the experimental results should be succinct but in sufficient detail to allow the experiments to be analyzed and interpreted by an independent reader. If necessary, subheadings may be used for an orderly presentation. All Figures and Tables should be referred to in the text in order, including those in the Supplementary Data.

Discussion: The data should be interpreted concisely without repeating material already presented in the Results section. Speculation is permissible, but it must be well-founded, and discussion of the wider implications of the findings is encouraged. Conclusions derived from the study should be included in this section.

Acknowledgments: All funding sources (including grant identification) should be credited in the Acknowledgments section. Authors should also describe the role of the study sponsor(s), if any, in study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication. If the funding source had no such involvement, the authors should so state.

In addition, people who contributed to the work but who do not meet the criteria for authors should be listed along with their contributions.

References: References should be numbered in the order in which they appear in the text. Citing of unpublished results, personal communications, conference abstracts, and theses in the reference list is not recommended but these sources may be mentioned in the text. In the reference list, cite the names of all authors when there are fifteen or fewer authors; if there are sixteen or more authors, list the first three followed by *et al*. Names of journals should be abbreviated in the style used in PubMed. Authors are responsible for the accuracy of the references. The EndNote Style of *Drug Discoveries & Therapeutics* could be downloaded at **EndNote** (*https://www.ddtjournal.com/examples/Drug_Discoveries_Therapeutics.ens*).

Examples are given below:

Example 1 (Sample journal reference):

Nakata M, Tang W. Japan-China Joint Medical Workshop on Drug Discoveries and Therapeutics 2008: The need of Asian pharmaceutical researchers' cooperation. Drug Discov Ther. 2008; 2:262-263.

Example 2 (Sample journal reference with more than 15 authors):

Darby S, Hill D, Auvinen A, *et al*. Radon in homes and risk of lung cancer: Collaborative analysis of individual data from 13 European case-control studies. BMJ. 2005; 330:223.

Example 3 (Sample book reference):

Shalev AY. Post-traumatic stress disorder: Diagnosis, history and life course. In: Post-traumatic Stress Disorder, Diagnosis, Management and Treatment (Nutt DJ, Davidson JR, Zohar J, eds.). Martin Dunitz, London, UK, 2000; pp. 1-15.

Example 4 (Sample web page reference):

World Health Organization. The World Health Report 2008 – primary health care: Now more than ever. *https://apps.who.int/iris/handle/10665/43949* (accessed September 23, 2022).

Tables: All tables should be prepared in Microsoft Word or Excel and should be arranged at the end of the manuscript after the References section. Please note that tables should not in image format. All tables should have a concise title and should be numbered consecutively with Arabic numerals. If necessary, additional information should be given below the table.

Figure Legend: The figure legend should be typed on a separate page of the main manuscript and should include a short title and explanation. The legend should be concise but comprehensive and should be understood without referring to the text. Symbols used in figures must be explained. Any individually labeled figure parts or panels (A, B, *etc*.) should be specifically described by part name within the legend.

Figure Preparation: All figures should be clear and cited in numerical order in the text. Figures must fit a one- or two-column format on the journal page: 8.3 cm (3.3 in.) wide for a single column, 17.3 cm (6.8 in.) wide for a double column; maximum height: 24.0 cm (9.5 in.). Please make sure that artwork files are in an acceptable format (TIFF or JPEG) at minimum resolution (600 dpi for illustrations, graphs, and annotated artwork, and 300 dpi for micrographs and photographs). Please provide all figures as separate files. Please note that low-resolution images are one of the leading causes of article resubmission and schedule delays.

Units and Symbols: Units and symbols conforming to the International System of Units (SI) should be used for physicochemical quantities. Solidus notation (*e.g*. mg/kg, mg/mL, mol/mm^2/min) should be used. Please refer to the SI Guide *www.bipm.org/en/si/* for standard units.

Supplemental data: Supplemental data might be useful for supporting and enhancing your scientific research and *Drug Discoveries & Therapeutics* accepts the submission of these materials which will be only published online alongside the electronic version of your article. Supplemental files (figures, tables, and other text materials) should be prepared according to the above guidelines, numbered in Arabic numerals (*e.g*., Figure S1, Figure S2, and Table S1, Table S2) and referred to in the text. All figures and tables should have titles and legends. All figure legends, tables and supplemental text materials should be placed at the end of the paper. Please note all of these supplemental data should be provided at the time of initial submission and note that the editors reserve the right to limit the size

and length of Supplemental Data.

7. Online Submission

Manuscripts should be submitted to *Drug Discoveries & Therapeutics* online at *https://www.ddtjournal.com/login*. Receipt of your manuscripts submitted online will be acknowledged by an e-mail from Editorial Office containing a reference number, which should be used in all future communications. If for any reason you are unable to submit a file online, please contact the Editorial Office by e-mail at *office@ddtjournal.com*

8. Accepted Manuscripts

Page Charge: Page charges will be levied on all manuscripts accepted for publication in *Drug Discoveries & Therapeutics* (Original Articles / Brief Reports / Reviews / Policy Forum / Communications: $140 per page for black white pages, $340 per page for color pages; News / Letters: a total cost of $600). Under exceptional circumstances, the author(s) may apply to the editorial office for a waiver of the publication charges by stating the reason in the Cover Letter when the manuscript online.

Misconduct: *Drug Discoveries & Therapeutics* takes seriously all allegations of potential misconduct and adhere to the ICMJE Guideline (*https://icmje.org/recommendations*) and COPE Guideline (*https://publicationethics.org/files/Code_of_conduct_for_journal_editors.pdf*). In cases of suspected research or publication misconduct, it may be necessary for the Editor or Publisher to contact and share submission details with third parties including authors' institutions and ethics committees. The corrections, retractions, or editorial expressions of concern will be performed in line with above guidelines.

(As of December 2022)

Drug Discoveries & Therapeutics
Editorial and Head Office
Pearl City Koishikawa 603,
2-4-5 Kasuga, Bunkyo-ku,
Tokyo 112-0003, Japan.
E-mail: office@ddtjournal.com