Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Comparative Effectiveness and Safety of Ustekinumab versus Anti–TNF Therapy in Crohn Disease: A Systematic Review

César Augusto Rojas Landa
cesar.landa@uabc.edu.mx

Universidad Autónoma de Baja California
Edwin Garcia Vasquez
Universidad de San Martín de Porres
Cesar Ayala Ugarte
Universidad Autónoma de Guadalajara
César Emmanuel Ibarra Hernández
Universidad de Guadalajara
Jesica Oseguera Rodríguez
Universidad Nacional Autónoma de México
Roberto Kevin Galan Orozco
Universidad de Guadalajara
Manuel Garza Castillo
Universidad de Monterrey
Emmanuel Amaro Hernandez
Tecnológico de Monterrey
Rocio Jaqueline Ramos Garcia
Universidad Anáhuac de Cancún
Lennin Fernando Lachino García
Universidad Michoacana de San Nicolás de Hidalgo
Miriam Alejandra Altamirano Herrera
Universidad Autónoma de Guadalajara
Alberto Ayala Aguilar
Desert Valley Hospital
Juan José Robalino Cerón
Universidad Técnica de Ambato
Mauricio Martín Fuentes Ampié
Universidad Cardenal Miguel Obando Bravo
Gabriel Ramírez Torres
Tecnológico de Monterrey
Benjamín Alejandro Rojas Landa
Hospital Regional de Especialidades Número 30, Instituto Mexicano del Seguro Social

Systematic Review

Keywords: Crohn disease, ustekinumab, anti-TNF, biologic therapy, treatment sequencing

Posted Date: August 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10360452/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Purpose:

The optimal sequencing of biologic therapies in Crohn disease remains uncertain, particularly when comparing anti–tumor necrosis factor (anti-TNF) agents and ustekinumab (UST). We performed a systematic review to evaluate their comparative effectiveness and safety.

Methods:

We conducted a systematic review following PRISMA 2020 guidelines and registered the protocol in PROSPERO (ID: 1057121). PubMed, ScienceDirect, and Cochrane Library were searched from September 17, 2024, to February 7, 2025. We included randomized and observational studies comparing UST and anti-TNF therapies in adult patients with Crohn disease. Outcomes included clinical response, remission, objective outcomes, and safety. Findings were synthesized narratively due to heterogeneity.

Results:

Ten studies were included, comprising retrospective cohorts and one randomized controlled trial. Early induction outcomes favored anti-TNF therapy in some studies, particularly for clinical response. However, clinical remission and objective outcomes were generally comparable across therapies. In biologic-experienced patients, UST demonstrated higher treatment persistence and lower discontinuation rates. Safety outcomes were similar across groups.

Conclusions:

UST and anti-TNF therapies show broadly comparable effectiveness and safety in Crohn disease. Anti-TNF agents may provide advantages in early response, while UST may offer greater durability in biologic-experienced patients. Individualized biologic selection and treatment sequencing strategies are supported by these findings.

INTRODUCTION

Crohn disease (CD) is a chronic inflammatory bowel disorder categorized by transmural inflammation that may involve any portion of the gastrointestinal tract, spanning from the mouth to the perianal area[1 2]. Perianal complications, including abscesses and fistulizing disease, represent severe manifestations associated with increased morbidity and surgical burden[3]. This condition is recognized for its progressive course and association with multiple complications which underscores the need for effective therapies[1 4 5].

The introduction of biological therapies has fundamentally transformed the therapeutic landscape for CD. Historically, anti-tumor necrosis factor-alpha (anti-TNF-α) antibodies (e.g., infliximab, adalimumab, certolizumab) have served as first-line advanced therapies, demonstrating significant efficacy in inducing remission and improving patient quality of life. However, a proportion of patients experience therapeutic failure, either due to lack of initial response or loss of response over time, requiring alternative biological therapy[6].

In this context, monoclonal antibody ustekinumab (UST), which targets the shared p40 subunit of interleukin-12 and interleukin-23, has emerged as a key alternative. Long-term extension studies have demonstrated sustained efficacy and safety of ustekinumab in Crohn disease[7], while additional clinical evidence has also supported the effectiveness of biologic therapies in specific phenotypes such as small-bowel stricturing disease[8]. Studies have validated its use in CD patients, particularly those who have failed previous anti-TNF-α therapies[9]. Its safety and effectiveness have also been evaluated in specific populations, including elderly patients with Crohn disease[10].

Given this scenario, there is a need for further investigation into the pharmacotherapeutic maintenance treatment of CD. This systematic review compares the efficacy and safety of ustekinumab and anti-TNF agents in Crohn disease, focusing on clinical response, remission, and safety outcomes to inform treatment selection.

Methods

This systematic review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines[11]. The protocol was registered in the PROSPERO database (ID: 1057121). The methodological approach was additionally informed by established frameworks for conducting systematic reviews, including structured processes for study identification, selection, and synthesis[12].

Artificial intelligence-assisted tools were used for language refinement and editorial support. All scientific interpretation, manuscript preparation, and final approval were performed by the authors.

Data Sources and Search Strategy

We conducted a literature search from 17 September 2024 through 7 February 2025 on PubMed, ScienceDirect, and the Cochrane Library. The search strategy combined Medical Subject Headings (MeSH) and free-text terms related to Crohn disease, ustekinumab, and anti-TNF agents. The search syntax included combinations such as:

(("Crohn Disease"[MeSH Terms] OR "Crohn's disease"[Title/Abstract] OR "inflammatory bowel disease"[Title/Abstract] OR "IBD"[Title/Abstract]) AND ("ustekinumab"[MeSH Terms] OR "ustekinumab"[Title/Abstract] OR "anti-IL12/23"[Title/Abstract]) AND ("anti-TNF"[Title/Abstract] OR "TNF inhibitors"[Title/Abstract] OR "tumor necrosis factor inhibitors"[MeSH Terms] OR "infliximab"[Title/Abstract] OR "adalimumab"[Title/Abstract] OR "certolizumab"[Title/Abstract]) AND ("randomized controlled trial"[pt] OR "controlled clinical trial"[pt] OR "clinical trial"[tw] OR random*[tw] OR "cohort study"[tw] OR "prospective studies"[MeSH Terms] OR "comparative study"[pt] OR "evaluation studies"[pt])).

No language restrictions were applied. Only studies conducted in humans were included. Filters were applied when available according to database specifications.

Selection Criteria

Types of participants

This review included adult patients (≥ 18 years old) with a confirmed diagnosis of CD who had received treatment with UST and/or anti-TNF agents. Studies were excluded if they involved pediatric populations, patients receiving other biologic therapies, or those with concomitant diseases that could interfere with outcome interpretation.

Types of intervention

This systematic review included studies that directly compared UST with anti-TNF agents—such as infliximab, adalimumab, or certolizumab—in adult patients with CD.

We included studies in which UST or anti-TNF agents were administered either as induction or maintenance therapy. Eligible studies enrolled patients with varying treatment histories, including both biologic-naïve individuals and those with prior exposure to biologics (e.g., anti-TNFs) or immunomodulators (e.g., azathioprine, methotrexate). To ensure comparability, we only included studies where both treatment arms (UST and anti-TNF) were clearly defined and evaluated under similar clinical conditions.

Studies were eligible if they assessed:

- - Patients receiving UST or anti-TNF as first-line biologic therapy, or
- - Patients who had previously failed or discontinued a biologic or immunomodulator, and subsequently initiated treatment with UST or another anti-TNF agent.

We excluded:

- - Studies evaluating only one of the two drugs without a direct comparator.
- - Studies comparing UST or anti-TNF agents with other biologics (e.g., vedolizumab, risankizumab).
- - Studies involving combination biologic therapy or treatment regimens that did not allow for isolated comparison.

Both intravenous (IV) and subcutaneous (SC) routes of administration were accepted, as long as they conformed to recognized induction and maintenance protocols as defined by the original studies.

Types of studies

This review included high-quality studies such as randomized controlled trials (RCTs), cohort studies, and case-control studies. We excluded review articles, meta-analyses, redundant publications, studies with insufficient data, case reports, case series, conference abstracts, non–peer-reviewed articles.

Type of outcomes

The primary outcomes of this review were the efficacy of UST and anti-TNF agents in the induction of clinical remission in adult patients with CD, as well as their effectiveness in clinical response and objective response. The latter included the evaluation of endoscopic remission, understood as the improvement of mucosal appearance, biochemical remission, endoscopic healing, among other parameters. On the other hand, the safety of these treatments was assessed in terms of the occurrence of adverse events, such as infections or infusion-related reactions. In addition, aspects related to treatment tolerability were considered as secondary outcomes.

Selection of Studies, data extraction, and screening.

Two independent reviewers (J.O.R. and M.G.C.) screened titles and abstracts using Rayyan software. Potentially eligible studies underwent full-text assessment by two additional independent reviewers (C.A.U. and M.A.A.H.). Disagreements at each stage were resolved by discussion and consensus with a third reviewer (C.A.R.L.). Data extraction was independently performed by C.E.I.H., R.K.G.O., E.A.H., R.J.R.G., L.F.L.G., A.A.A., J.J.R.C., and M.M.F.A. Extracted data were verified for accuracy by E.G.V. and C.A.R.L. Studies were included according to the predefined eligibility criteria established in the study protocol.

Data Evaluation: Assessment of Risk of Bias

The methodological quality of the included studies was assessed according to the *Cochrane Handbook for Systematic Reviews of Interventions*. The Newcastle-Ottawa Scale (NOS) was used to evaluate observational studies. Two independent reviewers (E.G.V. and C.A.U.) assessed the risk of bias, and any disagreements were resolved through discussion with a third reviewer (C.A.R.L.). Studies were classified according to the recommendations of the Cochrane Handbook and NOS guidance. Because substantial clinical and methodological heterogeneity was identified across the included studies, a quantitative meta-analysis was not performed, and findings were synthesized narratively.

Results

A total of 815 records were identified through three major databases: ScienceDirect (n = 513), PubMed (n = 252), and Cochrane (n = 50). No records were removed prior to screening. During the screening phase, 789 records were excluded based on title and abstract review, leaving 26 reports for full-text retrieval. Of these, one report could not be retrieved and 19 were excluded due to unsuitable study design (n = 4), population (n = 2), or comparator/drug studied (n = 13). As a result, 6 studies from databases met the inclusion criteria and were included in the review. Additionally, 76 records were identified through citation searching. After full-text assessment, 72 reports were excluded due to wrong comparator/drug (n = 42), wrong study design (n = 15), wrong population (n = 9), or other reasons (n = 6). Consequently, 4 additional studies were included from this source. In total, 10 studies fulfilled the eligibility criteria and were incorporated into the final review (Fig. 1).

Risk of bias assessment included nine retrospective cohort studies using the Newcastle-Ottawa Scale (Fig. 2). All studies were of good quality, with overall scores ranging from 7 to 9. Seven studies (77.7%) achieved the maximum score in the selection domain, reflecting adequate representativeness and definition of cohorts. In the comparability domain, most studies scored 2 points, although two studies (22.2%) scored only 1 point, indicating limited adjustment for potential confounders. Regarding the outcome/exposure domain, three studies (33.3%) achieved the highest score (3 points), while the remainder scored 2 points. In addition, one randomized controlled trial (RCT) included in the analysis was evaluated using the Cochrane RoB 2.0 tool (Fig. 3). This trial demonstrated a low risk of bias across all domains, with no concerns regarding randomization, deviations from intended interventions, missing outcome data, outcome measurement, or selective reporting.

Baseline characteristics of included studies

The included studies were in their majority retrospective cohorts, with some prospective cohorts one multinational RCT. Sample sizes ranged from 40 patients to 2758 participants. The mean age generally fell in the third and fourth decade of life, with follow-up periods ranging from 16 weeks to 40 months. CD duration ranged from 2.1 ± 0.9 years in Asian cohorts to 11.8 ± 11.8 years in European populations. Ileocolonic localization (L3) predominated, followed by isolated ileal (L1) and colonic (L2) involvement, while upper gastrointestinal disease (L4) was rare. In terms of behavior, non-stricturing/non-penetrating (B1) predominated, while B2 and B3 reached up to 30% in some anti-TNF groups.

In terms of treatment, IFX was usually administered IV at 5 mg/kg every eight weeks, while ADA was given with an induction of 160/80 mg SC followed by 40 mg every two weeks. UST was typically initiated with a weight-based IV induction (260–520 mg) and maintained with 90 mg SC every 8 or 12 weeks. Some studies allowed optimization strategies such as increasing anti-TNF doses or shortening intervals, whereas others shortened UST maintenance from 12 to 8 weeks in selected cases. These regimens reflect real-world practice and highlight the heterogeneity in CD management across diverse populations (Table 1).

Clinical response and remission

Early induction showed heterogeneous clinical response, with isolated advantages for anti-TNF. During weeks 4–16 Ahmed et al. favored ADA over UST (aOR 2.40) (23). In the propensity score–matched cohort by Rivière et al., at 3 months the effect favored anti-TNF over UST (aOR for UST compared with anti-TNF 0.30)[18]. By contrast, estimates were largely non-significant in Bokemeyer et al. (weeks 14–16)[24], He et al. (Crohn disease Activity Index "CDAI" at weeks 12 and 24[20], and the week-52 RCT by Sands et al[25]. (Table 2).

Clinical remission was mostly comparable overall, with selective advantages for anti-TNF. Rivière et al. favored ADA over UST at 3 months (aOR 2.26) and 12 months (aOR 2.62)[18]. Ahmed et al. also favored ADA during induction (aOR 2.35). No meaningful differences were observed in Bokemeyer et al. at end of induction[23], Moens et al. at weeks 26–52 (Harvey–Bradshaw Index "HBI"-based)[21], Huang et al. at weeks 14–26 and 44–56 (22), Na et al. across biologic-naïve and experienced groups[19], or Sands et al. at week 52[25] (Table 2).

Objective response and remission

Objective response was largely comparable across studies. Moens et al. showed a trend toward higher endoscopic response with ADA than with UST (aOR 2.24; p = 0.07)[21]. He et al. reported similar week-52 objective response using intestinal ultrasound plus endoscopy[20]. Consistently, Huang et al. found no differences in endoscopic or transmural response at weeks 14–26 and 44–56[22], and Sands et al. reported similar endoscopic response at week 52[25] (Table 2).

For objective remission, endoscopic remission was higher with ADA in Moens et al. (aOR 2.73)[21], while Sands et al. showed comparable endoscopic remission between ADA and UST at week 52[25]. In Huang et al., transmural remission was similar between IFX and UST at weeks 14–26 (17.2% and 16.4%) and weeks 44–56 (34.1% and 32.4%). In the same study, biochemical remission favored IFX at weeks 14–26 (aOR 1.79) and became comparable by weeks 44–56[22] (Table 2).

Treatment safety

Treatment safety outcomes were generally comparable between UST and anti-TNF therapies, with most studies reporting non-significant differences ($p > 0.05$) across evaluated endpoints. With regard to persistence and discontinuation, the results depended on the biological background. Na et al. showed lower discontinuation with UST in bio-experienced patients ($p = 0.004$) and a non-significant trend in bio-naïve patients[19], a pattern consistent with the RCT by Sands et al.,[25] which observed fewer interruptions with UST, and with the cohort by Eriksson et al. which described higher interruption rates with ADA. In contrast, Moens et al. and Rivière et al. found no differences[21 18]. Regarding adverse events, the overall rates were similar when reported, and discontinuations due to adverse events tended to be numerically lower with UST in several cohorts, although without cross-sectional statistical consistency. A detailed summary of safety outcomes, including adverse events, treatment discontinuation, and persistence across included studies, is presented in Table 3.

Table 1
Characteristics of eligible studies .

Author	Country	Study Design	Sample size	Age (years)	Follow up	Duration of illness	Disease localization	Disease behavior	History of Previous Abdominal Surgery related to CD
Eriksson et al. ‡ (2023)[16]	Sweden	Retrospective Cohort	anti-TNF: 156; UST: 156	anti-TNF: 41.5 ± 15.9; UST: 41.9 ± 14.4	40 mo (IQR: 23–52)	NR	NR	NR	NR
Shiga et al. ± (2022)[17]	Japan	Retrospective Cohort	anti-TNF: 1379; UST: 1379	anti-TNF: 37.6 ± 15.1; UST: 37.4 ± 14.9	2 y	UST: 3.4 y ± 5.9; anti-TNF: 3.3 y ± 6.0	NR	NR	anti-TNF: 164 (11.9); UST: 174 (12.6)
Rivière et al. ‡ (2023)[18]	France, Canada	Retrospective Cohort	anti-TNF: 69; UST: 39	anti-TNF: 33 (IQR: 26–46); UST: 35 (IQR: 23–49)	40 mo (IQR: 23–52)	NR	**L1** anti-TNF: 20 (29); UST: 9 (23) **L2** anti-TNF: 34 (49); UST: 15 (38) **L3** anti-TNF: 15 (22); UST: 15 (38)	**B1** anti-TNF: 37 (54); UST: 29 (74) **B2** anti-TNF: 21 (30); UST: 4 (10) **B3** anti-TNF: 11 (16); UST: 6 (15)	anti-TNF: 18 (26); UST: 7 (18)
Na et al. ‡ (2024)[19]	South Korea	Retrospective cohort	Bio-naïve IFX: 45; ADA: 14; UST: 59 Bio-Experienced anti-TNF:13; UST: 35	Bio-naïve IFX: 34 ± 14; ADA: 29 ± 8; UST: 35 ± 12 Bio-Experienced anti-TNF: 31 ± 8; UST: 32 ± 9	Bio-naïve IFX: 3.6 y ± 2.3; ADA: 3.9 y ± 2.2; UST: 2.1 y ± 0.9 Bio-experienced UST: 2.3 y ± 1.2; anti-TNF: 2.2 y ± 1.2	Bio-naïve IFX: 5.6 y ± 5.0; ADA: 5.3 y ± 5.3; UST: 5.9 y ± 6.3 Bio-experienced anti-TNF: 7.6 y ± 4.8; UST: 7.8 y ± 5.4	Bio-naïve **L1** IFX: 16 (35.6); ADA: 5 (35.7); UST: 27 (45.8) **L2** IFX: 5 (11.1); ADA: 2 (14.3); UST: 3 (5.1) **L3** IFX: 24 (53.3); ADA: 7 (50); UST: 29 (49.2) Bio Experienced **L1** Anti-TNF: 5 (38.5); UST: 12 (34.2) **L2**	Bio-naïve **B1** IFX: 31 (68.9); ADA: 7 (50); UST: 37 (62.7) **B2** IFX: 8 (17.8); ADA: 5 (35.7); UST: 15 (25.4) **B3** IFX: 6 (13.3); ADA: 2 (14.3); UST: 7 (11.9) Bio Experienced **B1** Anti-TNF: 4 (30.8); UST: 20 (57.1) **B2**	Bio-naïve patients: IFX: 8 (17.8); ADA: 6 (42.9); UST: 12 (20.3) Bio Experienced anti-TNF: 5 (38.5); UST: 9 (25.7)

Author	Country	Study Design	Sample size	Age (years)	Follow up	Duration of illness	Disease localization	Disease behavior	History of Previous Abdominal Surgery related to CD
							Anti-TNF: 1 (7.7); UST: 1 (2.9) **L3** Anti-TNF: 7 (53.8); UST: 22 (62.9)	Anti-TNF: 7 (53.8); UST: 9 (25.8) **B3** Anti-TNF: 2 (15.4); UST: 6 (17.1)	
He et al. (2024)[20]	China	Retrospective cohort	IFX: 106; UST: 105	IFX: 33.6 ± 15.6; UST: 33.3 ± 13	52 wk	IFX: 6.30 y ± 5.21; UST: 5.96 y ± 5.27	**L1** IFX: 17 (16.04); UST: 18 (17.14) **L2** IFX: 20 (18.87); UST: 18 (17.14) **L3** IFX: 69 (65.09); UST: 69 (65.71) **L4** IFX: 8 (7.55); UST: 9 (8.57)	**B1** IFX: 30 (28.3); UST: 23 (21.9) **B2** IFX: 41 (38.68); UST: 48 (45.71) **B3** IFX: 12 (11.32); UST: 14 (13.33) **B2 + B3** IFX: 23 (21.7); UST: 20 (19.04)	IFX: 37 (34.91); UST: 46 (43.81)
Moens et al. ‡ (2022)[21]	Belgium	Retrospective cohort	ADA: 32; UST: 32	ADA: 37 (IQR: 26–49); UST: 45 (IQR: 27–61)	60 wk (IQR: 33–104)	ADA: 8 y (IQR: 1–21); UST: 4 y (IQR: 4–24)	**L1** ADA: 14 (44);UST: 16 (50) **L2** ADA: 3 (9); UST: 1 (3) **L3** ADA: 15 (47); UST: 15 (47) **L4** ADA: 0 (0); UST: 2 (6)	**B1** ADA: 16 (50); UST: 16 (50) **B2** ADA: 9 (28); UST: 10 (31) **B3** ADA: 7 (22); UST: 6 (19)	ADA: 16 (50); UST: 15 (47)
Huang et al. ‡ (2025)[22]	China	Prospective cohort	IFX: 283; UST:146	IFX: 28 (IQR 22.3–34); UST: 27 (IQR 23–37)	56 wk	IFX: 19.7 mo (IQR: 9.5–60.2); UST: 24 mo (IQR: 11.5–62)	**L1** IFX: 40 (16.8); UST: 23 (16.3) **L2** IFX: 30 (12.6), UST: 17 (12.1) **L3** IFX: 168 (70.6), UST: 101 (71.6) **L4** IFX: 35 (14.7); UST: 21 (14.9)	**B1** IFX: 135 (56.7); UST: 83 (58.9) **B2** IFX: 61 (25.6); UST: 34 (24.1) **B3** IFX: 42 (17.7); UST: 24 (17.0)	IFX: 23 (9.7); UST: 14 (9.9)

Author	Country	Study Design	Sample size	Age (years)	Follow up	Duration of illness	Disease localization	Disease behavior	History of Previous Abdominal Surgery related to CD
Sands et al. (2022)[25]	Multi-country	RCT	ADA:195; UST:191	ADA: 37.4 ± 12.99; UST: 37.0 ± 13.23	1 y	ADA: 5.8 y ± 7.09; UST: 5.4 y ± 8.36	**L1** ADA: 55(28); UST: 60(32) **L2** ADA: 34(17); UST: 26(14) **L3** ADA:1 03(53); UST: 102(54) **L4** ADA: 17(9); UST: 30(16)	**B3*** ADA: 20 (10); UST: 17 (9)	ADA: 37 (19); UST: 29 (15)
Ahmed et al. (2019)[23]	USA	Ambidirectional cohort	ADA: 97; UST: 66	ADA: 40 ± 13.6; UST: 42.5 ± 15.9	16 wk	ADA: 8.6 y ± 9.4; UST: 11.8 y ± 11.8	**L1** ADA: 19 (19.6); UST: 12 (18.2) **L3** ADA: 78 (80.4); UST: 54 (81.8)	**B1** UST 19 (28.8); ADA 48 (49.5) **B2 + B3** ADA 49 (50.5); UST 47 (71.2)	NR
Bokemeyer et al. † (2023)[24]	Germany	Prospective cohort	ADA: 160; IFX: 119; UST: 339	anti-TNF: 38.2 (IQR: 28.3–50.4); UST: 37.3 (IQR: 29.7–50.0)	3 y	anti-TNF: 9.4 y (IQR: 3.0-17.3); UST: 8.8 y (IQR: 3.1–16.4)	**L1** anti-TNF: 74 (26.6); UST: 95 (27.9) **L2** anti-TNF: 52 (18.6); UST: 62 (18.2) **L3** anti-TNF: 153 (54.8), UST: 183 (53.9) **L4** UST: 60 (17.8); Anti-TNF: 49 (17.4)	**B1** anti-TNF: 137 (49.2); UST: 171 (50.6) **B2** UST: 100 (29.5); anti-TNF: 75 (26.8) **B3** UST: 68 (20); anti-TNF: 67 (24)	anti-TNF: 96 (34.6); UST: 120 (35.5)

ADA, adalimumab; Anti- TNF, anti-tumor necrosis factor agents; CD, Crohn disease; CI, confidence interval; Id, Induction dose; IFX, infliximab; IV, intravenous; Md, Maintenance dose; mo, months; q1w, once a week; q4w, every 4 weeks; q8w, every 8 weeks; q12w, every 12 weeks; RCT, Randomized Controlled Trial; SC, subcutaneous; UST, Ustekinumab; wk, week; y, years.

‡ The presence of fistulas implies penetrating disease (B3). Data on perianal disease were not considered because the data do not distinguish whether those with perianal involvement are B1, B2, or B3.

† Data reported after adjustment by propensity score with inverse probability of treatment weighting.

± Data reported after propensity score matching.

Studies that did not perform matching show the original data reported.

Table 2
Efficacy

Author	Clinical response		Clinical remission		Objective response		Objective remission	
	Definition	Findings	Definition	Findings	Definition	Findings	Definition	Findings
Rivière et al. ‡ (2023)[18]	≥ 3-point HBI decrease from baseline or HBI < 4, without steroids, CD surgery, or treatment discontinuation	At 3 mo anti-TNF: 59 (86); UST: 25 (64) Adjusted OR: 0.30 (CI 95%: 0.12 to 0.77); p value: 0.01	HBI < 4; without steroids, CD-related surgery, or treatment discontinuation due to failure/intolerance	At 3 mo ADA vs UST Adjusted OR: 2.26; IC 95%: 1.01 to 5.08; p: 0.05 At 12 mo Adjusted OR 2.62; IC 95%:1.18 to 5.79; p: 0.02	NR	NR	NR	NR
Na et al. ‡ (2024)[19]	NR	NR	CDAI score < 150 at response assessment after induction therapy	Bio-naïve IFX: 39 (86.7); ADA: 10 (71.4); UST: 55 (93.2); p: 0.071 Bio-experienced anti-TNF:10 (83.3); UST: 27 (77.1); p: 1.00	NR	NR	NR	NR
He et al. (2024)[20]	CDAI score decrease > 100 points from baseline.	Wk 12: IFX: 49 (46.23); UST: 40 (38.10) p: 0.232 Wk 24: IFX: 58 (54.72); UST: 53 (50.48); p: 0.537	NR	NR	IUS response (≥ 2 mm reduction in BWT) + Endoscopic response (reduction in stenosis with increased luminal diameter compared to baseline)	Wk 52 IFX: 8/12 (66.67); UST: 16/21 (76.19); p: 0.69	NR	NR
Moens et al. ‡ (2022)[21]	NR	NR	HBI < 5; if unavailable, physician's global assessment was used	Wk 26 ADA vs UST Adjusted OR: 1.26 (95% CI: 0.36 to 4.51); p: 0.72 Wk 52 Adjusted OR: 1.58 (95% CI: 0.54 to 4.88); p: 0.4	Endoscopic response: ≥ 50% reduction in SES-CD from baseline or SES-CD ≤ 3	Wk 26–52 ADA vs UST Adjusted OR: 2.24 (95% CI: 0.94–5.71); p: 0.07	Endoscopic remission: SES-CD score < 3	Wk 26–52 ADA vs UST Adjusted OR 2.73 (95% CI: 1.12–7.36); p: 0.03

Author	Clinical response		Clinical remission		Objective response		Objective remission	
	Definition	Findings	Definition	Findings	Definition	Findings	Definition	Findings
Huang et al. ‡ (2025)[22]	NR	NR	CDAI score < 150	Wk 14–26 IFX: 172 (72.3); UST: 91 (64.5) Rate difference 7.7% (95% CI − 2.0 to 17.5; p: 0.115). Adjusted OR 1.56 (95% CI 0.95 to 2.56; p: 0.076) Wk 44–56 IFX: 174/233 (74.7); UST: 101/138 (73.2) Rate difference 1.5% (95% CI − 7.8 to 10.8; p: 0.751). Adjusted OR 1.06 (95% CI 0.65 to 1.74; p: 0.819)	Endoscopic response: > 50% decrease in SES-CD Transmural response: > 25% decrease in BWT in the most affected segment on IUS	Endoscopic response: Wk 14–26: IFX: 143(60.1); UST: 78 (55.3) Rate difference 4.8% (95% CI − 5.5 to 15.1; p: 0.363). Adjusted OR 1.25 (95% CI 0.80 to 1.95; p: 0.333) Wk 44–56 IFX: 125/212 (59.0); UST: 69/118 (58.5) Rate difference 0.5% (95% CI − 10.6 to 11.6; p: 0.931). Adjusted OR 1.12 (95% CI 0.73 to 1.74; p: 0.601). Transmural response: Wk 14–26: IFX: 74/180 (41.1); UST: 43/116 (37.1). Rate difference 4.0% (95% CI − 7.3 to 15.4; p: 0.487). Adjusted OR 1.19 (95% CI 0.72 to 1.98; p: 0.490). Wk 44–56 IFX: 87/170 (51.2); UST: 53/111 (47.8). Rate difference 3.4% (95% CI − 8.5 to 15.4; p: 0.574). Adjusted OR 1.10 (95% CI 0.68 to 1.79; p: 0.699).	Endoscopic remission: SES-CD ≤ 2 (or SES-CD = 0 for patients with baseline SES-CD of 3) Transmural remission: BWT ≤ 3.0 mm in all intestinal segments (≤ 4.0 mm in sigmoid colon), color Doppler signal ≤ 1, normal bowel wall stratification, and absence of inflammatory mesenteric fat. Biochemical remission: CRP level < 5 mg/L	Endoscopic remission: Wk 14–26: IFX: 87 (36.6); UST: 43 (30.5) Rate difference 6.1% (95% CI − 3.7 to 15.8; p: 0.230). Adjusted OR 1.31 (95% CI 0.81 to 2.12; p: 0.262). Wk 44–56 IFX: 90/212 (42.5); UST: 49/118 (41.5) Rate difference 0.9% (95% CI − 10.2 to 12.0; p: 0.870). Adjusted OR 1.12 (95% CI 0.71 to 1.76; p: 0.637). Transmural remission: Wk 14–26 IFX: 31/180 (17.2); UST: 19/116 (16.4). Rate difference 0.8% (95% CI − 7.9 to 9.5; p: 0.850). Adjusted OR 1.19 (95% CI 0.72 to 1.98; p: 0.490). Wk 44–56 IFX: 58/170 (34.1); UST: 36/111 (32.4). Rate difference 1.7% (95% CI − 9.6 to 12.9; p: 0.770). Adjusted OR 1.08 (95% CI 0.66 to 1.78; p: 0.755). Biochemical remission: Wk 14–26 IFX: 172 (72.3); UST: 87/140 (62.1). Rate difference 10.1% (95% CI 0.3–20.0; p: 0.041). Adjusted OR 1.79 (95% CI

Author	Clinical response		Clinical remission		Objective response		Objective remission	
	Definition	Findings	Definition	Findings	Definition	Findings	Definition	Findings
								1.11–2.90; p: 0.017). Wk 44–56 IFX: 142/224 (63.4); UST: 88/134 (65.7). Rate difference −2.3% (95% CI: − 12.5 to 7.9; p: 0.663). Adjusted OR 0.90 (95% CI 0.58–1.41; p: 0.646).
Ahmed et al. (2019)[23]	HBI score decrease ≥ 2 with max score ≤ 7 at response assessment	wk 4–16 ADA: 71 (73.2); UST: 33 (50) Adjusted OR: 2.40 (95% CI: 1.14 to 5.07); p: 0.02)	Clinical response (HBI decrease ≥ 2, max 7) plus HBI < 4 at follow-up	wk 4–16 ADA: 43/97 (44.3); UST: 18/65 (27.7) Adjusted OR: 2.35 (95% CI: 1.07 to 5.16); p: 0.0337	NR	NR	NR	NR
Sands et al. (2022)[25]	CDAI score decrease ≥ 100 from baseline or CDAI score < 150	Wk 52 ADA:129 (66); UST: 138 (72) Group difference: 6% (95% CI − 3 to 15); p: 0.18	CDAI score < 150	Wk 52 ADA: 119 (61); UST:124 (65) Group difference: 4% (95% CI: − 6 to14); p: 0.42	Endoscopic Response: SES-CD reduction ≥ 50% from baseline; or SES-CD ≤ 3; or SES-CD 0 if baseline SES-CD 3	Wk 52 ADA: 66/179 (37); UST: 75/179 (42) Difference 5% (95% CI: − 5 to15); p: 0.35	Endoscopic Remission: SES-CD ≤ 3; or SES-CD 0 if baseline SES-CD 3	Wk 52 ADA: 55/179 (31); UST: 51/179 (29) Difference −2% (95% CI: − 12 to7); p: 0.63
Bokemeyer et al. † (2023)[24]	≥ 3-point HBI reduction and/or remission at end of induction	wk 14–16 UST vs anti-TNF OR: 1.04 (IC 95%: 0.65 to 1.67), p:0.866	HBI ≤ 4 at end of induction	wk 14–16 anti-TNF: 176 (63.0); UST: 222(65.4) OR: 1.11 (IC 95%: 0.71 to 1.74); p: 0.647	NR	NR	NR	NR

ADA, adalimumab; anti-TNF, anti-tumor necrosis factor agents; BWT, Bowel wall thickness; CD, Crohn's disease; CDAI, Crohn's disease activity index; CD-PRO-2, Crohn disease Patient-Reported Outcomes-2; CI, Confidence Interval; CRP, C-reactive protein; HBI, Harvey-Bradshaw index; IFX, infliximab; IUS, intestinal ultrasound; mo, months; NR, Not Reported; OR, Odds ratio; p, p- value; SES-CD; Simple Endoscopic Score for CD; UST, Ustekinumab; vs, versus; wk, week.

† Data reported after adjustment by propensity score with inverse probability of treatment weighting.

‡ Data reported in the propensity score-matched cohort.

Studies that did not perform matching show the original data reported.

Table 3
Safety

Author	Adverse events	Hospitalization	Surgery	Discontinuation of the drug	
				Total	**Reasons**
Eriksson et al. ‡ (2023)[6]	NR	Anti-TNF: 47 (30); UST: 44 (28); p: 0.99	Anti-TNF: 12 (8); UST: 20 (13); p: 0.33	Anti-TNF: 103 (66.03); UST: 91 (58.3)	Due to loss of response Anti-TNF: 30 (29.1); UST: 26 (28.6) Due to adverse events Anti-TNF: 21 (20.4); UST: 11 (12.1) Due to other reasons Anti-TNF: 52 (50.5); UST: 54 (59.3)
Shiga et al. ‡ (2022)[17]	NR	Anti-TNF: 281 (20.4); UST: 273 (19.8); p: 0.85	Anti-TNF: 67 (4.86); UST: 55 (3.99); p: 0.46	Anti-TNF: 127 (9.2) at 1 y, 259 (18.8) at 2 y UST: 103 (7.5) at 1 y, 350 (25.4) at 2 y p: 0.46	NR
Rivière et al. ‡ (2023)[18]	NR	NR	NR	Anti-TNF vs UST: no difference; p: 0.29	NR
Na et al. ‡ (2024)[19]	NR	**Bio-naïve** IFX: 6 (13.3); ADA: 1 (7.1); UST: 8 (13.6); p: 0.800 **Bio-experienced** Anti-TNF: 4 (30.8); UST: 12 (34.3); p: 1.000	**Bio-naïve** IFX: 3 (6.7); ADA: 1 (7.1); UST: 1 (1.7); p: 0.390 **Bio-experienced** Anti-TNF: 2 (15.4); UST: 6 (17.1); p: > 0.05	**Bio-naïve** IFX: 6 (13.3); ADA: 2 (14.3); UST: 1 (1.7); p: 0.052 **Bio-experienced** Anti-TNF: 7 (53.8); UST: 4 (11.4); p: 0.004	NR
He et al. (2024)[20]	IFX: 9/108 (8.3); UST: 6/106 (5.7)	NR	IFX: 2/56 (3.57); UST: 3/63 (4.76); p: 0.99	NR	Due to adverse events IFX: 2/108 (1.89); UST: 1/106 (0.95)
Moens et al. ‡ (2022)[21]	Within 1 y ADA: 15 (46.88); UST: 17 (53.13) After 1 y ADA: 19 (59); UST: 18 (56)	NR	ADA: 2 (5); UST: 3 (9); p: 0.65	ADA: 15 (46.88); UST: 15 (46.88)	Due to loss of response ADA: 3 (9); UST: 3 (9); p: 1.00 Due to primary non-response ADA: 5 (16); UST: 8 (25); p: 0.54 Due to adverse events ADA: 7 (22); UST: 3 (9); p: 0.30 Due to loss to follow-up ADA 0 (0); UST: 1 (3); p: 1.00
Huang et al. ‡ (2025)[22]	NR	NR	IFX: 4 (1.4); UST: 5 (3.5)	IFX: 58 (24.4); UST: 29 (20.6); p: 0.372	Due to lack of response IFX: 51 (21.4); UST: 25 (17.7) Adverse drug-related events IFX: 5 (2.1); UST: 1 (0.7) Loss to follow-up

Author	Adverse events	Hospitalization	Surgery	Discontinuation of the drug	
				Total	Reasons
					IFX: 2 (0.8%); UST: 3 (2.1) Bowel surgery IFX: 4 (1.4%); UST: 5 (3.4)
Sands et al. (2022)[25]	ADA 152 (78); UST 153 (80)	NR	NR	ADA: 46 (24); UST: 29 (15)	Due to adverse events ADA: 21 (11); UST: 11 (6) Due to withdrawal of consent ADA: 10 (5); UST: 11 (6) Due to absence of improvement ADA: 10 (5); UST: 4 (2) Due to lost to follow-up ADA: 2 (1); UST: 2 (1) Due to pregnancy ADA: 1 (< 1); UST: 1 (< 1)
Bokemeyer et al. †[24]	NR	UST:16 (4.72); Anti-TNF:14 (5.02); p: > 0.05	NR	NR	NR

ADA, adalimumab; anti-TNF, anti-tumor necrosis factor agents; NR, Not Reported; p, p- value; UST, Ustekinumab; vs, versus; y, year.

† Data reported after adjustment by propensity score with inverse probability of treatment weighting.

‡ Data reported after propensity score matching.

Studies that did not perform matching show the original data reported.

Discussion

This systematic review summarizes the available head-to-head and quasi–head-to-head evidence comparing UST and anti-TNF agents in CD. Overall, UST and anti-TNF therapies showed broadly comparable effectiveness and safety across most clinical and objective outcomes, with differences that tended to depend on previous biologic exposure and the specific outcome assessed rather than indicating clear global superiority of one strategy.

In biologic-naïve populations, cohorts by Rivière et al., Moens et al. and Ahmed et al. generally reported higher rates of clinical response or remission with ADA or anti-TNF agents compared with UST[18 21 23]. By contrast, the SEAVUE trial found no significant differences between UST and ADA in clinical or endoscopic remission at week 52 in biologic-naïve CD[25]. Taken together, these data indicate that UST and anti-TNF agents are both valid first-line options in moderate-to-severe CD, and that modest discrepancies between cohorts are likely driven by differences in study design, dosing strategies, follow-up duration and patient selection rather than a consistent superiority signal of one agent over the other. This interpretation is in line with recent positioning papers and guidelines that still regard IFX and ADA as reference first-line biologics, while acknowledging UST as a reasonable alternative in a treat-to-target framework[6].

Among bio-experienced patients, especially those with prior anti-TNF failure, the pattern is different. Studies by Ahmed et al., Na et al., and Eriksson et al.[16 19 23], consistently suggest higher treatment persistence, lower discontinuation rates, and clinically meaningful responses with ustekinumab (UST) compared with infliximab (IFX) or adalimumab (ADA) in this setting[16 19 23]. These findings are supported by additional real-world evidence demonstrating improved treatment persistence with UST compared with anti-TNF agents in biologic-experienced populations[24 26]. These findings do not demonstrate consistent superiority of UST but suggest that, after anti-TNF failure, switching to a biologic with a different mechanism of action may provide greater durability, likely due to lower immunogenicity and distinct pharmacokinetic properties. Taken together, these findings support considering UST as a second-line option in refractory CD or in patients with prior biologic exposure. This is consistent with algorithms that prioritize anti-TNF as first-line therapy and place UST after exposure to or intolerance of these drugs[6 28].

The comparative interpretation of these findings is limited by heterogeneity in outcome definitions across studies. Clinical response and remission were defined using either the CDAI or the HBI. Although these scores are highly correlated and a 3-point change in HBI roughly corresponds to a 100-point change

in CDAI, they are not fully interchangeable and may classify individual patients differently. Similarly, objective outcomes were assessed using different endoscopic scores, various time points and, in some studies, without a clear definition of transmural response. This variability implies that small numerical differences between UST and anti-TNF arms should be interpreted with caution, as they may reflect the choice of measurement tool as much as the underlying treatment effect.

Biochemical outcomes add another layer of complexity. Biomarkers such as C-reactive protein and faecal calprotectin correlate with endoscopic activity and are central components of treat-to-target strategies, but they do not fully capture mucosal or transmural healing and can be influenced by extra-intestinal inflammation or comorbid conditions[27]. In several included cohorts, biochemical response or remission was not systematically paired with endoscopic or radiologic evaluation, limiting the ability to interpret biochemical normalisation as definitive evidence of superiority for either UST or anti-TNF agents. Furthermore, patient-level variability in disease burden, pharmacokinetics, immunogenicity and comorbidities can modulate the effectiveness of any biologic, independent of its class. For these reasons, biochemical endpoints in this review are best viewed as supporting elements within a composite target, rather than stand-alone determinants of comparative advantage.

The heterogeneity observed in study designs and outcome definitions also reflects external contextual factors. The STRIDE-II recommendations emphasise treat-to-target strategies aiming for clinical remission, endoscopic healing and, when feasible, transmural healing, but their implementation depends on local resources, access to endoscopy and cross-sectional imaging, and the availability of serial biomarker testing[28]. At the same time, ECCO guidelines and other international consensus documents often recommend TNF inhibitors as first-line advanced therapy in moderate-to-severe CD, while placing UST and other biologics as subsequent options after anti-TNF failure or in specific clinical scenarios[629]. Economic evaluations also influence these choices, with several models identifying IFX or ADA as more cost-effective in some health-care systems, while others support the cost-effectiveness of UST in particular contexts[28]. These external factors likely contribute to the diversity of treatment sequences, the predominance of certain comparators and the variability in targets chosen across the included studies.

Compared with previous systematic reviews and meta-analyses that primarily evaluated UST versus placebo or relied on indirect comparisons between biologic classes, our review focuses on direct or quasi-direct comparisons between UST and anti-TNF agents across multiple domains, including clinical, endoscopic, biochemical and treatment-persistence outcomes. Our results are broadly consistent with the meta-analysis by Chenna et al., which did not detect major differences between UST and ADA in clinical or endoscopic remission or steroid-free remission[27]. Furthermore, pivotal randomized clinical trial evidence has demonstrated the efficacy of ustekinumab for both induction and maintenance therapy in Crohn disease[29]. Together, these findings support a personalised rather than a “one-size-fits-all” approach to the choice between UST and anti-TNF therapy.

Additionally, methodological advances in systematic review reporting and data visualization, such as those proposed in PRISMA2020 extensions and associated tools, have improved transparency and reproducibility in evidence synthesis[30].

From a clinical perspective, the available evidence supports both UST and anti-TNF agents as appropriate options for biologic-naïve patients with moderate-to-severe CD, with no consistent global differences in efficacy or safety when used as first-line biologics. In many settings, anti-TNF agents remain a common first-line choice on the basis of longer clinical experience, cost and guideline recommendations, whereas UST appears particularly attractive after anti-TNF failure or intolerance and in clinical scenarios where lower immunogenicity, specific safety considerations or convenient maintenance dosing are prioritised. In practice, treatment decisions should integrate prior biologic exposure, disease location and behaviour, history of surgery, extra-intestinal manifestations, comorbidities, patient preferences and economic and organisational factors, in line with treat-to-target principles.

This review also has important limitations that mirror the underlying evidence base. Most of the included studies are observational, with inherent risks of residual confounding and selection bias, and only one large randomised head-to-head trial directly compares UST and ADA in biologic-naïve CD. Sample sizes are modest in several cohorts, definitions of response and remission are variable, and the timing and type of objective assessments are not standardised. These factors precluded robust quantitative synthesis across all outcomes and limit the generalisability of pooled estimates. Publication bias and selective reporting cannot be excluded, particularly for under-reported outcomes such as transmural healing, health-related quality of life and economic endpoints. Consequently, our conclusions should be interpreted as hypothesis-generating and as a framework to guide clinical reasoning rather than as definitive evidence of superiority for one strategy.

Future research should prioritise adequately powered, head-to-head randomised controlled trials and large comparative cohorts that use harmonised definitions of clinical, endoscopic, transmural and biochemical outcomes and pre-specify analyses by prior biologic exposure, surgery and disease phenotype. Integrating therapeutic drug monitoring, predictive biomarkers and patient-reported outcomes into treat-to-target strategies may help refine sequencing algorithms and improve long-term disease control, while robust cost-effectiveness and implementation data are needed to ensure that advances in biologic therapy translate into equitable access for patients with CD across different health-care systems.

Conclusion

Evidence shows that UST and anti-TNF agents have broadly comparable effectiveness and safety across most clinical and objective outcomes in CD, with low rates of surgery and hospitalization reflecting similar safety profiles. For biologic-naïve patients, some studies have shown a selective advantage for anti-TNF (particularly ADA) in early clinical response or remission. In contrast, in biologic-experienced patients, UST consistently demonstrated greater treatment persistence and lower discontinuation rates compared with anti-TNF, suggesting a clinical advantage for switching to UST in this context.

Declarations

Competing Interests

The authors have no competing interests to declare that are relevant to the content of this article.

Ethics Approval

This systematic review was conducted in accordance with PRISMA 2020 guidelines, and registered in PROSPERO (ID: 1057121). Ethical approval and informed consent were not required because this study analyzed previously published data without direct patient involvement.

Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Funding

No funding was received for conducting this study.

Author Contribution

C.A.R.L., E.G.V., and C.A.U. contributed equally to this work and share first authorship. C.A.R.L., E.G.V., and C.A.U. conceived and designed the study. J.O.R., M.G.C., C.A.U., M.A.A.H., C.E.I.H., R.K.G.O., E.A.H., R.J.R.G., L.F.L.G., A.A.A., J.J.R.C., and M.M.F.A. participated in the literature search, study screening, full-text assessment, data extraction, and interpretation of the findings. C.A.R.L. drafted the initial manuscript. G.R.T. and B.A.R.L. contributed to the preparation and revision of the tables and figures and critically revised the manuscript. E.G.V. and C.A.U. critically reviewed the manuscript for important intellectual content. All authors reviewed, edited, approved the final version of the manuscript, and agree to be accountable for all aspects of the work.

Acknowledgement

The authors would like to thank Oliver Alejandro Rojas Landa, Christian Orlando Wertz Bush, Raymundo Carrasco Campos, Fátima Guadalupe Ramírez Ibarra, Andrea Natalia Bruno Bernabé, Tania Alejandra Ramírez Ramírez, Kenneth Miguel León Cruz, Miriam Nohemí Martínez Brownell, for their technical guidance and support regarding the software and digital tools used during the conduct of this systematic review.

Data Availability

No new datasets were generated during this study. All data analyzed were obtained from previously published studies included in this systematic review. All relevant data are contained within the article and its supplementary materials. Additional information regarding study selection and extracted data is available from the corresponding author upon reasonable request.

References

1. Baumgart DC, Sandborn WJ. Crohn disease. Lancet. 2012;380(9853):1590–605. 10.1016/S0140-6736(12)60026-9.
2. Roda G, Ng SC, Kotze PG, et al. Crohn disease. Nat Rev Dis Primers. 2020;6:22. 10.1038/s41572-020-0156-2.
3. Makowiec F, Jehle C, Becker HD, et al. Perianal abscess in Crohn disease. Langenbecks Arch Chir. 1993;378(3):164–9. 10.1007/BF00181360.
4. Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century. Lancet. 2017;390:2769–78. 10.1016/S0140-6736(17)32448-0.
5. Dorn-Rasmussen M, Lo B, Zhao M, et al. Incidence and prevalence of inflammatory bowel disease in Denmark 1980–2017. J Crohns Colitis. 2023;17:259–68. 10.1093/ecco-jcc/jjac162.
6. Lichtenstein GR, Loftus EV Jr, Isaacs KL, et al. ACG clinical guideline: management of Crohn disease in adults. Am J Gastroenterol. 2018;113:481–517. 10.1038/ajg.2018.27.
7. Sandborn WJ, Rebuck R, Wang Y, et al. Five-year efficacy and safety of ustekinumab in Crohn disease: the IM-UNITI trial. Clin Gastroenterol Hepatol. 2021;20:578–90. 10.1016/j.cgh.2020.05.059.
8. Bouhnik Y, Carbonnel F, Laharie D, et al. Adalimumab effectiveness in small-bowel Crohn disease stricture: the CREOLE study. Gut. 2018;67:53–60. 10.1136/gutjnl-2016-312579.
9. Barkai LJ, Gonczi L, Balogh F, et al. Ustekinumab efficacy and safety over three years in Crohn disease. Sci Rep. 2024;14:1–8. 10.1038/s41598-023-50134-2.
10. Gebeyehu GG, Broglio G, Liu E, et al. Safety and effectiveness of ustekinumab vs anti-TNF therapy in elderly patients with Crohn disease. Inflamm Bowel Dis. 2024. 10.1093/ibd/izae174.

11. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. 10.1136/bmj.n71.
12. Khan KS, Kunz R, Kleijnen J, Antes G. Five steps to conducting a systematic review. J R Soc Med. 2003;96:118–21. 10.1258/jrsm.96.3.118.
13. Ouzzani M, Hammady H, Fedorowicz Z, et al. Rayyan: a web and mobile app for systematic reviews. Syst Rev. 2016;5:210. 10.1186/s13643-016-0384-4.
14. Higgins JPT, Thomas J, Chandler J, et al. editors. Cochrane handbook for systematic reviews of interventions. Version 6.4. Cochrane; 2024.
15. Wells GA, Shea B, O'Connell D et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute. Accessed May 14, 2026. Newcastle-Ottawa Scale.
16. Eriksson C, Söderling J, Karlqvist S, et al. Ustekinumab vs anti-TNF therapy as second-line treatment in Crohn disease. Dig Dis Sci. 2023;68:3119–28. 10.1007/s10620-023-07897-2.
17. Shiga H, Tarasawa K, Moroi R, et al. Long-term effectiveness of ustekinumab vs anti-TNF therapy in Crohn disease. J Gastroenterol Hepatol. 2022;37:2105–12. 10.1111/jgh.15992.
18. Rivière P, Kanters C, Pellet G, et al. Ustekinumab vs anti-TNF as first-line biologic therapy in Crohn disease. Inflamm Bowel Dis. 2023;29:923–31. 10.1093/ibd/izac167.
19. Na JE, Park YE, Park J, et al. Real-world outcomes comparing ustekinumab, infliximab, and adalimumab in Crohn disease. BMC Gastroenterol. 2024;24. 10.1186/s12876-024-03403-8.
20. He X, Wang Y, Sun J, et al. Ustekinumab vs infliximab in Crohn disease with stenosis. Ther Adv Gastroenterol. 2024;17. 10.1177/17562848241290663.
21. Moens A, Alsoud D, Verstockt B, et al. Adalimumab vs ustekinumab as first-line biologic therapy in Crohn disease. Eur J Gastroenterol Hepatol. 2022;34:1015–20. 10.1097/MEG.0000000000002411.
22. Huang Z, Diao N, Guo Q, et al. Infliximab vs ustekinumab for endoscopic and transmural remission in Crohn disease. Clin Gastroenterol Hepatol. 2025. 10.1016/j.cgh.2024.12.040.
23. Ahmed Z, Venkata K, Zhang N, et al. Ustekinumab vs adalimumab for induction therapy in Crohn disease. Gastroenterol Res. 2019;12:245–51. 10.14740/gr1194.
24. Bokemeyer B, Plachta-Danielzik S, di Giuseppe R, et al. RUN-CD study: ustekinumab vs anti-TNF therapy in Crohn disease. Inflamm Bowel Dis. 2023;29:1741–50. 10.1093/ibd/izac271.
25. Sands BE, Irving PM, Hoops T, et al. Ustekinumab vs adalimumab for induction and maintenance therapy in Crohn disease: the SEAVUE study. Lancet. 2022;399:2200–11. 10.1016/S0140-6736(22)00688-2.
26. Rayer C, Nachury M, Arnaud B, et al. Ustekinumab, vedolizumab, or second anti-TNF therapy after failure of first anti-TNF therapy in Crohn disease. BMC Gastroenterol. 2022;22. 10.1186/s12876-022-02583-5.
27. Chenna VSH, Nagi TK, Suarez ZK, et al. Ustekinumab vs adalimumab in Crohn disease: a systematic review and meta-analysis. Cureus. 2023. 10.7759/cureus.38277.
28. Turner D, Ricciuto A, Lewis A, et al. STRIDE-II: an update on the Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) initiative of the International Organization for the Study of IBD: determining therapeutic goals for treat-to-target strategies in inflammatory bowel disease. Gastroenterology. 2021;160(5):1570–83. 10.1053/j.gastro.2020.12.031.
29. Feagan BG, Sandborn WJ, Gasink C, et al. Ustekinumab as induction and maintenance therapy for Crohn disease. N Engl J Med. 2016;375:1946–60. 10.1056/NEJMoa1602773.
30. Haddaway NR, Page MJ, Pritchard CC, et al. PRISMA2020: an R package and Shiny app for producing PRISMA 2020-compliant flow diagrams. Campbell Syst Rev. 2022;18:e1230. 10.1002/cl2.1230.

Figures

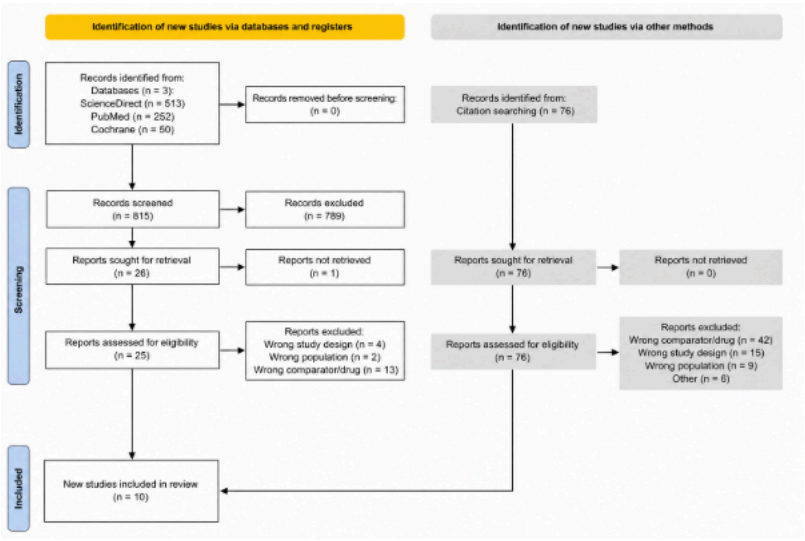


Figure 1

Study selection flow diagram following PRISMA methodology. This stepwise illustration provides transparency in our filtering process, which guided the final study selection.

Author, year	Selection (max 4 stars)	Comparability (max 2 stars)	Outcome/Exposure (max 3 stars)	Total (max 9 stars)	Subjective Evaluation
Eriksson C, et al. (2023)[16]	4	2	2	8	Good quality
Shiga H, et al. (2022)[17]	4	2	2	8	Good quality
Rivière P, et al. (2023)[18]	4	2	3	8	Good quality
Na JE, et al. (2024)[19]	4	2	2	8	Good quality
He X, et al. (2024)[20]	4	1	3	7	Good quality
Moens A, et al. (2022)[21]	4	2	3	9	Good quality
Huang Z, et al. (2025)[22]	3	2	2	7	Good quality
Ahmed Z, et al. (2019)[23]	4	1	2	7	Good quality
Bokemeyer B, et al. (2023)[24]	4	2	2	8	Good quality

Figure 2

Risk of bias assessment using NOS for Cohort studies. This figure summarized the risk of bias across domains for each cohort. The evaluation was executed using the Newcastle-Ottawa Scale.

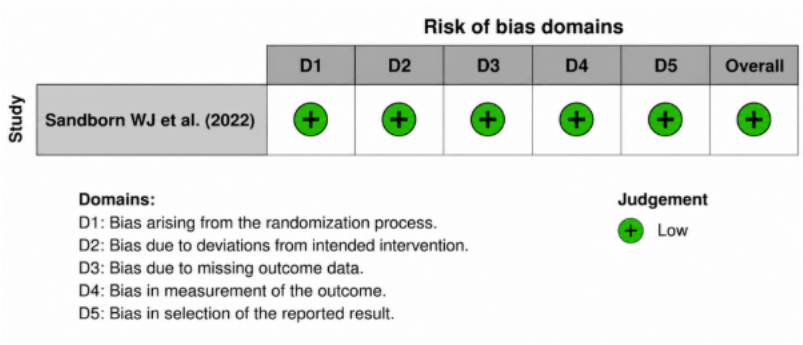


Figure 3

Risk of bias assessment across domains for RCT. The figure summarizes the risk of bias across domains for the trial evaluated. The evaluation was executed using the Cochrane RoB 2.0 tool.