*Annals of Gastroenterology* (2026) **39,** 1-11

# Evaluating the role of capsule endoscopy for non-responsive, refractory and complicated celiac disease: a systematic review and meta-analysis

**Ian Io Lei[a,b,c]*, Rinda Naresh[c]*, Anastasios Koulaouzidis[d,e], Ramesh P. Arasaradnam[a,b,f]**

University Hospital of Coventry and Warwickshire, Coventry UK; University of Warwick, Coventry, UK; Royal Berkshire Hospital, Reading, UK; University of Southern Denmark, Odense, Denmark; Odense University Hospital, Denmark; University of Leicester, UK

## Abstract

**Background** Small-bowel capsule endoscopy (CE) is used in selected adults with nonresponsive, refractory, or complicated celiac disease (CD), but its diagnostic yield (DY), clinical impact, and safety remain uncertain.

**Methods** We searched Medline, EMBASE, CENTRAL, and PubMed from January 2000 to November 30, 2025, for studies of CE in adults with established CD and persistent symptoms, suspected refractory disease, or suspected complications. Two reviewers independently selected studies, extracted data, and assessed risk of bias using ROBINS-E; certainty was assessed using GRADE. Random-effects meta-analyses using generalized linear mixed models estimated pooled DY and secondary outcomes. Meta-regression and subgroup analyses were considered exploratory.

**Results** Seventeen studies, including 994 adults, were eligible. The pooled DY of CE in complex CD was 82% (95% confidence interval [CI] 67-91%), with substantial heterogeneity ($I^2$=82.5%). DY was highest in refractory CD cohorts. Pooled DY was 3% for ulcerative jejunitis and 3% for small-bowel malignancy. Capsule retention was uncommon. CE findings were associated with management changes in several studies; however, this outcome was reported in only 9 studies, and definitions varied. Egger's test and trim-and-fill analysis suggested possible small-study effects, reducing the adjusted DY estimate to 79%. The pooled conversion to device-assisted enteroscopy was estimated at 18% (95%CI 9-30%).

**Conclusions** CE may be a valuable second-line, noninvasive triage tool in selected high-risk patients with complex CD, particularly refractory CD. However, the available evidence remains of very low certainty and should be interpreted cautiously, given the substantial methodological and clinical heterogeneity across studies.

**Keywords** Capsule endoscopy, small bowel, diagnostic yield, refractory celiac disease, complicated celiac disease

*Ann Gastroenterol 2026; 39 (XX): 1-11*

Correspondence to: Ian Io Lei, Clifford Bridge Rd, Binley, Coventry CV2 2DX, UK, e-mail: brian.lei@uhcw.nhs.uk

Received 11 January 2026; accepted 23 May 2026; published online 9 July 2026

DOI: https://doi.org/10.20524/aog.2026.1092



## Introduction

Celiac disease (CD) is a chronic, immune-mediated enteropathy triggered by gluten exposure in genetically susceptible individuals [1,2]. Although adherence to a strict gluten-free diet (GFD) remains the cornerstone of treatment, up to 30% of adults continue to experience ongoing symptoms despite reported adherence [3,4]. These cases may represent non-responsive CD (NRCD) or refractory CD (RCD), which are associated with nutritional deficiencies, impaired quality of life, and an elevated risk of lymphoproliferative malignancy [5-7].

NRCD is characterized by persistent symptoms, abnormal serology or ongoing mucosal abnormalities after 6-12 months of a strict GFD. In most cases, NRCD reflects inadvertent gluten exposure or alternative gastrointestinal (GI) pathology (e.g., irritable bowel syndrome), rather than true treatment failure [3,4]. In contrast, RCD is defined by persistent or recurrent malabsorptive symptoms with ongoing villous atrophy, despite strict GFD for ≥12 months, after exclusion of alternative causes of enteropathy [8,9]. RCD is further subdivided into type I, characterized by a normal intraepithelial lymphocyte phenotype, and type II, defined by an aberrant clonal intraepithelial lymphocyte population, which carries a substantially poorer prognosis [8,9]. Complicated CD (CCD) encompasses structural and malignant sequelae, including ulcerative jejunitis, strictures and enteropathy-associated T-cell lymphoma (EATL); these typically arise in longstanding NRCD or RCD and carry high morbidity and mortality [8,9]. Early recognition and precise differentiation of these disease states is critical for guiding appropriate further management.

For clinicians managing patients who have persistent symptoms despite confirmed dietary adherence, determining the optimal diagnostic pathway remains challenging. Traditional evaluation relies on repeated esophagogastroduodenoscopy with duodenal biopsies; however, this is invasive, prone to sampling error, and limited to the proximal small bowel (SB) [10]. SB capsule endoscopy (CE) provides a minimally invasive assessment of the entire SB, and is more sensitive than conventional endoscopy for detecting villous atrophy and complications in established CD [10,11]. According to the European Society of Gastrointestinal Endoscopy (ESGE) guidelines, CE is recommended in suspected NRCD or RCD as a second-line test, after serology, biopsies and dietary review, to assess villous atrophy and identify complications [12]. Cross-sectional imaging and device-assisted enteroscopy (DAE) are generally used for the targeted characterization or biopsy of lesions identified on CE [10,11,13].

Despite these recommendations, the evidence supporting CE in NRCD or CCD remains fragmented and of low certainty. Existing meta-analyses have primarily examined CE at initial diagnosis, rather than its diagnostic yield (DY) in high-risk NRCD, RCD or CCD cohorts [14,15]. Evidence regarding the clinical impact of CE, specifically how findings influence subsequent investigations or management, remains limited [10]. We therefore conducted a systematic review and meta-analysis to evaluate the DY of CE in NRCD, RCD and CCD, identify factors associated with diagnostic performance, and assess procedure-related complications.

## Materials and methods

This study protocol was designed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines (see supplemental table 1) [16]. The protocol was registered with the PROSPERO International Register of Systematic Reviews (Registration ID: CRD420251243243). No amendments were made to the protocol following registration. This study aimed to systematically evaluate the DY, clinical impact and safety of CE in adult patients with NRCD, RCD, or CCD.

### Eligibility criteria

We included full-text observational studies (retrospective or prospective) that evaluated the use of CE in adults with established CD presenting with NRCD, RCD or CCD. All types of CE were included, and all referral indications were considered, including symptomatic patients, asymptomatic patients undergoing evaluation, positive serology or previously confirmed RCD. Conference abstracts were excluded because of insufficient methodological detail and/or limited extractable data. Review articles, systematic reviews, editorials, study protocols, case reports and studies with <10 participants were excluded.

### Information sources

We conducted a comprehensive literature search across the following electronic databases: Medline (Ovid), EMBASE (Ovid), Cochrane Central Register of Controlled Trials (CENTRAL), and PubMed. The search strategy incorporated controlled vocabulary terms (e.g., MeSH terms) and free-text keywords relevant to CD and CE. The complete search strategy for each database is available in the Supplementary Materials (Supplementary Table 2), in accordance with the PRISMA-S (PRISMA-Search) guidelines. The search was limited to studies published between 1 January 2000 and 30 November 2025, and only English language publications were included. To ensure comprehensiveness, we also conducted a manual review of reference lists of full-text articles and relevant systematic reviews.

[a]Institute of Precision Diagnostics and Translational Medicine, University Hospital of Coventry and Warwickshire, Coventry, UK (Ian Io Lei, Ramesh P. Arasaradnam); [b]Warwick Medical School, University of Warwick, Coventry, UK (Ian Io Lei, Ramesh P. Arasaradnam); [c]Department of Gastroenterology, Royal Berkshire Hospital, Reading, UK (Ian Io Lei, Rinda Naresh); [d]Department of Clinical Research, University of Southern Denmark, Odense, Denmark (Anastasios Koulaouzidis); [e]Department of Surgery, Odense University Hospital, Denmark (Anastasios Koulaouzidis); [f]Leicester Cancer Center, University of Leicester, UK (Ramesh P. Arasaradnam)

*These authors contributed equally to this work as joint first authors

Conflict of Interest: IIL, RN and RPA report no competing interests. AK received research support from Given Imaging Ltd. (via the ESGE) and IntroMedic; has served as a consultant for Jinshan Ltd.; and has received consultancy fees from Jinshan, lecture honoraria from Covidien/Medtronic, and educational travel support or honoraria from Jinshan, Dr Falk Pharma UK, Ferring, Aquilant, and Almirall. He has also served on advisory boards for Tillotts, Ankon, Jinshan, and Dr Falk Pharma UK, and is co-founder and director of iCERV Ltd

### Study selection

All records were imported into Rayyan (Rayyan Systems Inc.), a web-based platform designed for systematic review screening. Duplicate entries were identified and removed. Title and abstract screening, followed by full-text eligibility assessment, were independently performed by 2 reviewers (IIL and RN). Disagreements were resolved through discussion and consensus.

### Inclusion criteria

- Participants had a confirmed diagnosis of CD established by serology and/or duodenal histology
- The study included adults with established CD undergoing CE for any clinical indication
- The sample size was ≥10 participants
- Any CE system was permitted, including gastric, SB, colon, or panenteric capsules, irrespective of manufacturer
- The study did not need CD as a primary or secondary endpoint, provided that relevant CE outcomes were extractable

### Exclusion criteria

- Pediatric studies (<18 years of age) or non-human studies
- Studies with suspected but unconfirmed CD
- Studies focusing exclusively on obscure GI bleeding or non-CD indications
- Studies using CE primarily for the initial diagnosis of CD, rather than for evaluation, surveillance or assessment of complications

### Data compilation

The final set of studies was reviewed, and data extraction was independently performed by IL and RN. Extracted variables included patient demographics, disease duration and indications for CE. Clinical data included the number of patients with NRCD, RCD or CCD, subsequent DAE, elevated biomarkers, and symptomatic presentations. NRCD refers to persistent symptoms, abnormal serology or mucosal abnormalities, despite 6-12 months of a GFD [3,4]. RCD was defined by ongoing malabsorptive symptoms and villous atrophy, despite ≥12 months of GFD, and both Type 1 and 2 were included [8,9]. CCD encompassed structural or malignant complications arising from longstanding NRCD or RCD, including ulcerative jejunitis, strictures and EATL. For the purposes of this meta-analysis, complex CD (CxCD) was used as an operational umbrella term encompassing NRCD, RCD, and CCD; however, these entities are clinically distinct and should not be considered equivalent in terms of disease severity, risk profile or DY. Villous atrophy on CE was identified by macroscopic surrogate markers, including scalloping of mucosal folds, mosaic pattern, fissuring, nodularity, and absence of villi. DY was defined as the proportion of CE examinations identifying 1 or more clinically significant small-bowel abnormalities in patients with CxCD. The histological reference standard varied across studies, and the Marsh–Oberhuber classification based on duodenal biopsies was most commonly used, while flow cytometry for aberrant intraepithelial lymphocytes was used to subtype RCD in 3 studies [17-19]. Technical details included CE system type, completion rate, bowel preparation quality, use of patency capsule and adverse events. Diagnostic outcomes included overall and lesion-specific DY (ulcers, ulcerative jejunitis, Crohn's disease, malignancy), and whether CE changed diagnosis or altered patient management.

### Risk of bias

The risk of bias (RoB) for all non-randomized studies was evaluated using the ROBINS-E tool for non-randomized studies of exposures. We acknowledge that QUADAS-2 is the standard tool for diagnostic accuracy studies; however, given that our review focused on DY and downstream clinical impact, rather than sensitivity and specificity, and included observational studies without a uniform reference standard, we selected ROBINS-E as a more appropriate framework to assess bias relevant to an exposure–outcome paradigm. ROBINS-E assesses potential bias arising from confounding, participant selection, exposure classification, deviations from intended exposure, missing data, outcome measurement and selective reporting [20]. Two reviewers independently assessed each domain and assigned ratings of low, moderate, serious or critical risk, with disagreements resolved by consensus or third-reviewer adjudication; overall study ratings reflected the highest-risk domain. The certainty of evidence for each outcome was assessed using the GRADE framework, considering RoB, inconsistency, indirectness, imprecision and publication bias [21].

### Statistical analysis

We performed random-effects meta-analyses using a generalized linear mixed model to estimate pooled DYs for active RCD, as well as secondary outcomes, including ulcerative jejunitis, impact on management, conversion to DAE, lymphoma, and non-lymphomatous small-bowel tumors. Heterogeneity was assessed using the $I^2$ statistic and the Cochran's Q test. To explore potential sources of heterogeneity, we performed univariate meta-regression using the "rma" function in the "metafor" package [22]. Effect sizes were calculated as logit-transformed proportions using the "escalc" function. Prespecified moderators included mean age, ulcer and tumor DY, CE completion rate, proportion with positive celiac serology (tTG), proportion with established RCD, and symptom burden. Subgroup and sensitivity analyses were performed based on study-level characteristics and RoB.

All analyses were conducted using R (version 4.3.2) [23] with the "meta" [24], "metafor" and "dplyr" [25] packages.

## Results

### Literature search and study selection

A systematic search conducted on 30 November 2025 identified 574 records across 4 databases. Following duplicate removal and title/abstract screening, 27 articles underwent full-text review, from which 17 studies met the inclusion criteria for the systematic review and meta-analysis. The reasons for the exclusion of the remaining 11 studies are detailed in Fig. 1, with the most common being inclusion of suspected rather than confirmed CD, and conference-only abstracts. One additional study focusing on chronic diarrhea in patients with established CD was identified through reference checking. Of the included studies, 6 were retrospective and the remainder prospective, contributing a total of 994 patients to the meta-analysis (Supplementary Table 3). The mean age across cohorts ranged from 45.9-55.8 years (Table 1).

Across the 17 included studies, a wide variety of positive findings encompassed: (i) macroscopic markers of villous atrophy, including scalloping of mucosal folds, mosaic pattern, fissuring, nodularity, and flattening or absence of villi; (ii) mucosal inflammation, including erosions, erythema and edema; (iii) ulcerative lesions, including aphthous ulcers and ulcerative jejunitis or jejunoileitis; (iv) stricturing disease; and (v) suspected small-bowel malignancy, including mass lesions and circumferential thickening. These features are consistent with the established CE criteria for CD described in the ESGE CE guidelines [26] and the European Society for the Study of Celiac Disease's 2025 updated diagnostic guidelines [12]. In addition, the definitions and granularity of "management change" varied considerably across studies, ranging from broad categories to specific therapeutic interventions (Supplementary Table 4).

The RoB assessment was performed using the ROBINS-E tool, summarized in Supplementary Table 5. The majority of

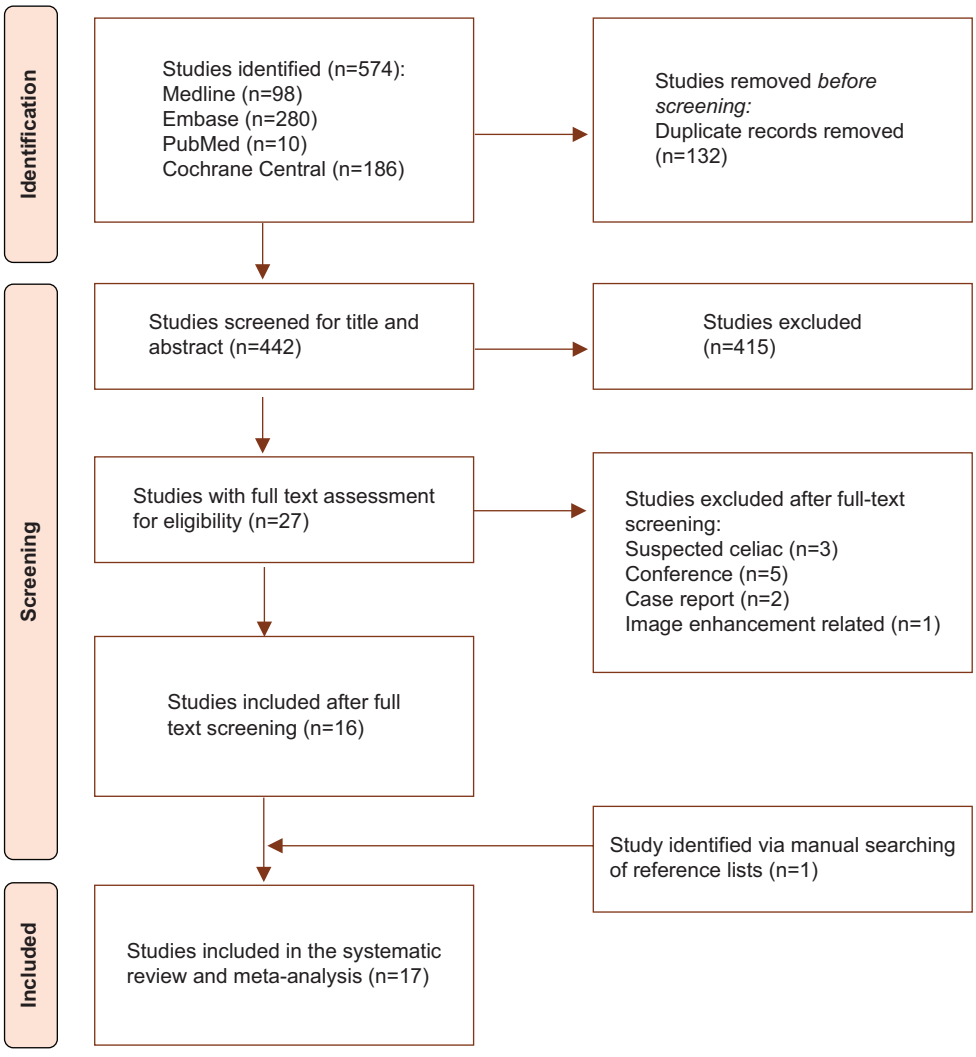



**Figure 1** PRISMA flowchart

**Table 1** Characteristics of included studies

| Author and Location [ref.] | Year | Study type | Mean/ median Age | Age range or SD | Bowel cleansing adequacy | CE completion rate | Male sex, proportion | Sample size included/enrolled (pediatric/adult) | CE type (manufacturer, model) |
|---|---|---|---|---|---|---|---|---|---|
| Daum *et al* (Germany)[33] | 2007 | Single-center, retrospective | 51.0 | (28-85) | 0.86 | 0.64 | 0.5 | 14 | M2A capsule |
| Maiden *et al*(United Kingdom) [34] | 2009 | Single-center, prospective | 57.0 | (18-82) | - | 0.84 | 0.26 | 19 | M2A capsule |
| Barret *et al* (France) [35] | 2012 | Single-center, retrospective | 47.0 | - | 0.53 | 0.62 | 0.32 | 38 | M2A capsule |
| Tomba *et al* (Italy) [36] | 2014 | Single-center, prospective | 49.0 | 4.6 | 1 | 0.96 | 0.19 | 52 | PillCam SB2 |
| Perez *et al* (Spain) [37] | 2018 | Multi-center, retrospective | 46.6 | 16.6 | 1 | 0.95 | 0.30 | 189 | PillCam SB3/Mirocam |
| Chetcuti Zammit *et al* (United Kingdom) [38] | 2019 | Single-center, prospective | 52.5 | 15.5 | - | 0.96 | 0.28 | 23 | PillCam SB3 |
| Branchi *et al.* (Italy) [11] | 2020 | Single-center, prospective | 51.2 | 16.6 | 0.88 | 1 | 0.08 | 13 | PillCam SB3 |
| Branchi *et al* (Italy) [11] | 2020 | Single-center, prospective | 51.2 | 16.6 | 1 | 0.85 | 0.08 | 16 | Capsocam |
| Chetcuti Zammit *et al* (United Kingdom) [4] | 2020 | Single-center, prospective | 54.3 | 15.7 |  | 0.93 | 0.3 | 100 | PillCam SB3 |
| Ferretti *et al* (Italy) [30] | 2022 | Single-center, prospective | 49.0 | 16 | 0.96 | 0.96 | 0.34 | 151 | PillCam SB3 |
| Katsinelos *et al* (Greece) [39] | 2011 | Multi-center, prospective | 45.9 | 16.7 | - | - | 0.49 | 10 | Not specified |
| Kurien *et al* (United Kingdom) [40] | 2013 | Single-center, prospective | 56.0 | (22-83) | - | - | 0.30 | 69 | PillCam SB or PillCam SB2 |
| Shiha *et al* (United Kingdom) [17] | 2025 | Multi-center, retrospective | 53.0 | 35-65 | - | 0.91 | 0.36 | 146 | PillCam SB3 |
| Elli *et al* (Italy) [18] | 2023 | Multi-center, retrospective | 52.5 | (50-62) | - | - | 0.30 | 66 | M2A, M2A Plus, PillCam SB2, PillCam SB3, and PillCam Crohn's capsule. |
| Chetcuti Zammit *et al* (United Kingdom) [19] | 2020 | Single-center, prospective | 49.7 | 16.9 | - | 0.97 | 0.40 | 35 | PillCam SB2/3 |
| Atlas *et al* (USA) [41] | 2011 | Single-center, prospective | 55.8 | 13 | - | 0.76 | 0.24 | 42 | M2A capsule |
| Foerster *et al* (Germany) [42] | 2021 | Single-center, retrospective | 48.0 | - | 0.91 | 1 | 0.36 | 11 | PillCam Crohn's capsule |

*CE, capsule endoscopy; SD, standard deviation*

studies were observational and predominantly single center, with small sample sizes and substantial heterogeneity in study design, patient selection and outcome definitions. RoB was rated as serious in 5 studies, particularly for selection bias, unclear indications, and non-consecutive or highly preselected cohorts. Reader blinding was inconsistently applied, and reference standards varied substantially across studies, with many lacking systematic histological confirmations or relying only on biopsies of high-risk lesions.

As summarized in Supplementary Table 6, the certainty of the evidence, as assessed using GRADE, across the included studies was predominantly very low. This was driven by small observational designs, indirectness and substantial imprecision. Only 2 studies achieved a low-certainty rating, with secondary outcomes similarly limited by inconsistent reference standards, unblinded outcome assessment, and sparse reporting of clinical impact or adverse events.

## Overall results

The pooled DY of CE in CxCD was 82% (95%CI 67-91%), with substantial heterogeneity ($I^2$=82.5%). This estimate should be interpreted in the context of heterogeneous definitions of positive CE findings across studies. Subgroup analysis identified disease phenotype as the primary determinant of yield, with the highest DY observed in RCD (99%, 95%CI 82-100%; $I^2$=0) (Fig. 2). The pooled DY for CCD further highlighted this pattern, with a markedly higher yield in the RCD subgroup (18%, 95%CI 4-56%; $I^2$=86.3%) compared with NRCD (4%, 95%CI 2-10%; $I^2$=0%), indicating that complications are concentrated in RCD populations (Fig. 3). Accordingly, pooled estimates should not be interpreted uniformly across disease phenotypes. CE findings were associated with changes in clinical management in a substantial proportion of patients (Supplementary Fig. 1), although this outcome was highly heterogeneous and is presented descriptively. The pooled conversion rate to DAE was 18% (95%CI 9-30%; $I^2$=88.8%), supporting a role for CE as a triage investigation, with only a minority requiring further invasive assessment. For specific findings, the pooled DY was 9% (95%CI 5-16%; $I^2$=70.2%) for ulcers, 3% (95%CI 1-7%; $I^2$=64.3%) for confirmed ulcerative jejunitis, and 3% (95%CI 1-6%) for malignancy (Supplementary Fig. 2,3). Capsule retention was uncommon (0.05%, 95%CI 0.02-0.14%). Completion rates ranged from 62-100%, and bowel preparation adequacy from 53-100%. PillCam (Medtronic, USA) was the most frequently used system (8 studies), followed by M2A (Given Imaging, Israel) (5 studies), panenteric PillCam Crohn's capsule (2 studies), and MiroCam (IntroMedic, South Korea) (1 study).

Publication bias was evaluated using funnel plot inspection, Egger's regression test, and the trimandfill method. The funnel plot visualization suggested mild asymmetry, which was confirmed statistically by Egger's regression test (t=2.19, df=15, P=0.0446), suggesting potential small-study effects or publication bias. Application of the trim-and-fill method imputed 4 potentially missing studies on the left side of the funnel plot, shifting the pooled estimate downward to 79% (standard error 34%; P=0.0186) (Supplementary Fig. 4). Although the adjusted effect remained statistically significant, the attenuation suggests that the observed DY may be modestly inflated in the published literature.

### Meta-regression

Meta-regression analyses were undertaken to explore covariates that might influence the DY of CE in CxCD. The exploratory univariate model was examined, with the results summarized in Supplementary Table 7. Two factors demonstrated significant associations with higher DY for positive CE findings: the proportion of patients with known RCD and the proportion of patients whose clinical management was altered following CE. Studies enrolling a greater proportion of patients with established RCD consistently reported higher DYs.

A dedicated meta-regression evaluating the DY of malignancy was also performed. Known RCD again emerged as a significant predictor of higher malignancy detection (P=0.006), as shown in the bubble plot (Supplementary Fig. 5). Similarly, a higher reported impact of CE on clinical decision-making was strongly associated with more malignant DY (P=0.004). Sensitivity analyses using leave-one-out procedures showed stable effect estimates across all studies. No individual study exerted disproportionate influence on the pooled results (Supplementary Fig. 6).

### Subgroup analysis

A subgroup analysis was undertaken using covariates identified as significant in the meta-regression. The relationship between DY and the impact of CE on subsequent management was first examined. Studies were stratified by the proportion of patients whose management changed following CE (high >55%, moderate 30-55%, low <30%). Only 9 eligible studies reported this outcome. The high-impact group demonstrated a nearly maximal pooled DY of 99% (95%CI 20-100%; $I^2$=0%), with progressively lower yields in the moderate- and low-impact groups; the between-group difference was significant (P<0.001) (Fig. 4). This gradient suggests a direct association between DY and downstream changes in management.

A further subgroup analysis assessed whether the proportion of patients with known RCD influenced tumor-specific DY. Studies without established RCD reported a pooled malignant DY of 2% (95%CI 1-4%; $I^2$=0%), whereas studies including patients with known RCD showed a higher yield of 4% (95%CI 2-10%; $I^2$=69.5%) (Supplementary Fig. 3). This pattern reflects the recognized increased risk of EATL and ulcerative jejunoileitis in established RCD, particularly type 2. Subgroup analyses by age, disease duration, study design and CE type revealed no significant differences between strata.

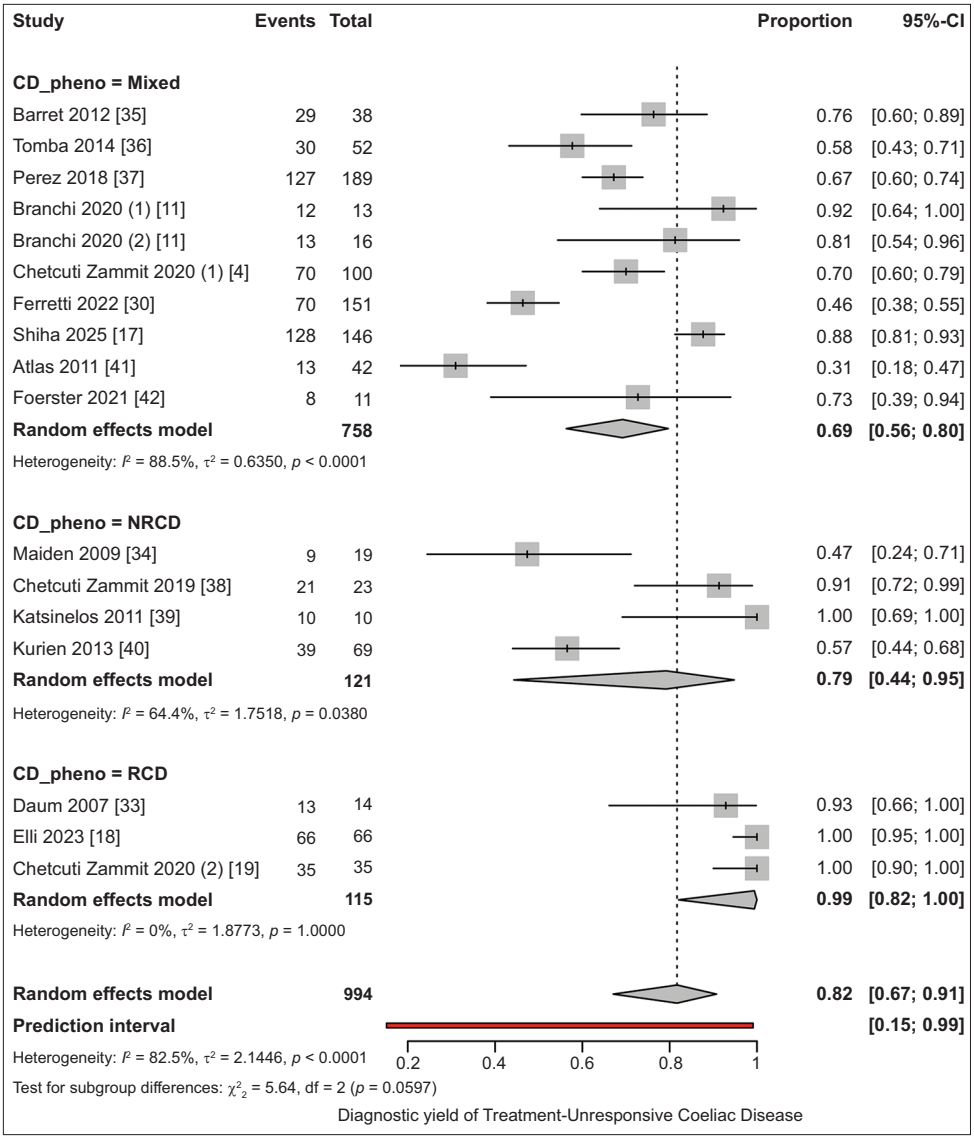



**Figure 2** Pooled diagnostic yield of capsule endoscopy in CxCD, stratified by disease phenotype. Forest plot demonstrating the overall pooled diagnostic yield of 82% (95%CI 67-91%) with substantial heterogeneity ($I^2$=82.5%). Subgroup analysis shows marked variation by phenotype, with the highest yield observed in RCD (99%, 95%CI 82-100%) compared with NRCD (79%, 95%CI 44-95%) and mixed (69%, 95%CI 56-80%) cohorts. The prediction interval (15-99%) highlights considerable between-study variability, indicating that diagnostic yield should be interpreted in the context of the underlying disease phenotype

*CxCD, complex celiac disease; NRCD, non-responsive celiac disease; RCD, refractory celiac disease; CI, confidence interval; $I^2$, Higgins heterogeneity statistic; $\tau^2$, between-study variance; df, degrees of freedom; $\chi^2$, chi-squared statistic*

## Discussion

CD, particularly when associated with persistent villous atrophy, carries an increased risk of complications and mortality. Current ESGE guidelines emphasize clinical review, serology, and repeat duodenal biopsies for surveillance, reserving CE and DAE for the uncommon scenario of suspected NRCD or RCD. These targeted modalities are therefore used selectively in a high-risk subgroup. However, no previous systematic review and meta-analysis has specifically evaluated the DY of CE in this context, its influence on subsequent management, or the covariates that may affect DY. Furthermore, the certainty of evidence supporting the use of CE in patients with persistent symptoms despite strict dietary adherence, seronegative CD, or suspected RCD or CCD remains limited [12].

This work presents the first systematic review and meta-analysis to characterize the DY of CE in NRCD, RCD and CCD. Across studies, the pooled DY was 82% (95%CI 67-91%), a figure comparable to the DY of 80% (95%CI 74-87%) reported for CE within 48 h of overt GI bleeding [27], and broadly consistent with the DY range of 55-68% described in Crohn's disease [28,29]. These findings demonstrate that CE has a high DY for detecting clinically important abnormalities in

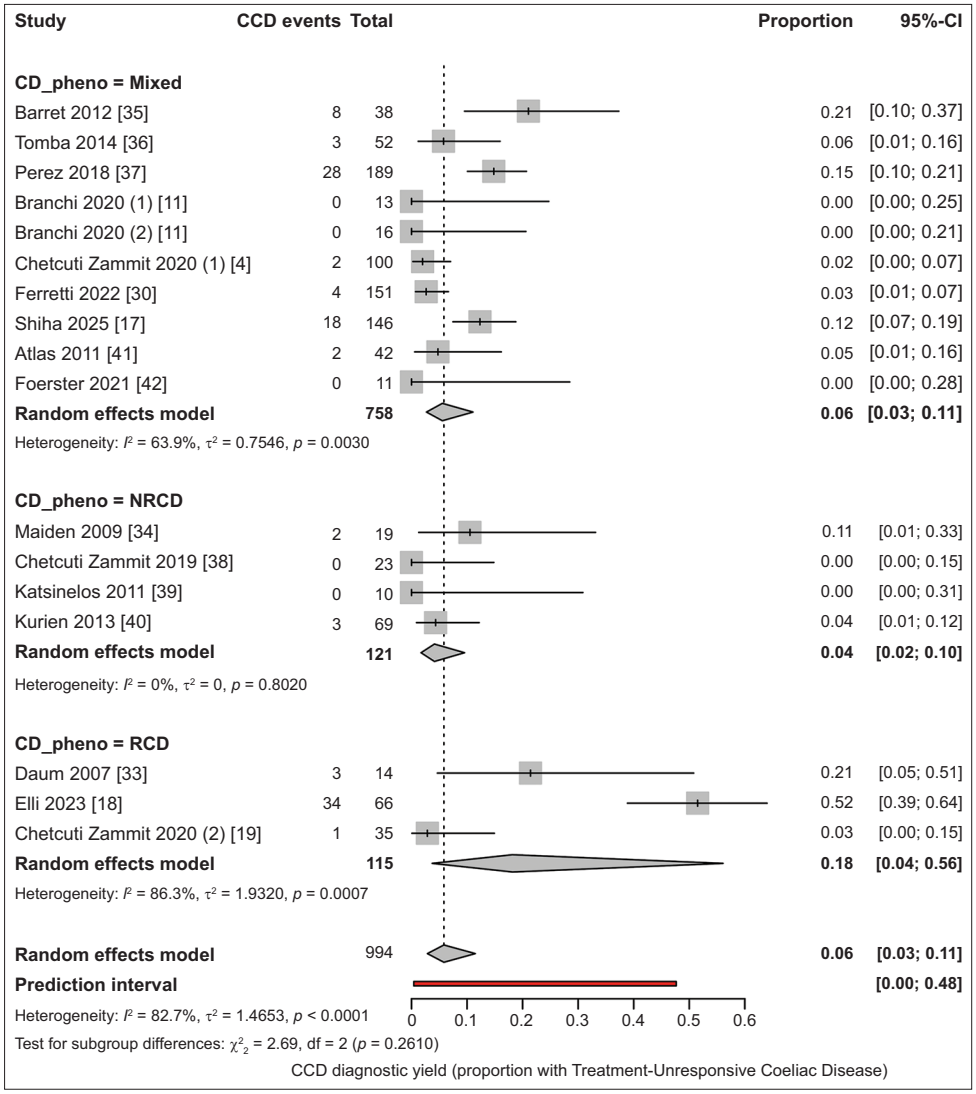



**Figure 3** Pooled diagnostic yield of CCD, stratified by disease phenotype within the CxCD cohort. Forest plot showing an overall pooled CCD diagnostic yield of 6% (95%CI 3-11%) with substantial heterogeneity ($I^2$=82.7%). Subgroup analysis demonstrates higher yield in RCD (18%, 95%CI 4-56%) compared with NRCD (4%, 95%CI 2-10%) and mixed (6%, 95%CI 3-11%) cohorts, highlighting the concentration of complications in RCD populations. The wide prediction interval (0-48%) reflects considerable between-study variability

*CCD, complicated celiac disease; CxCD, complex celiac disease; NRCD, non-responsive celiac disease; RCD, refractory celiac disease; CI, confidence interval; $I^2$, Higgins heterogeneity statistic; $\tau^2$, between-study variance; df, degrees of freedom; $\chi^2$, chi-squared statistic*

this population, including villous atrophy, erosions, ulcerative jejunoileitis and EATL. Diagnostic performance in this cohort mirrored that of DAE, which achieved an 83% yield (19/23 cases) in the sole published series [30]. The RCD subgroup exhibited a DY of 99% ($I^2$=0). In contrast, the NRCD subgroup had the lowest CCD DY, at 4% (95%CI 2-10%; $I^2$=0%), compared with 18% in the RCD group. These findings demonstrate the critical role of CE in the high-risk RCD cohort. Despite a high DY of 82% with CE, DAE was performed in only 18% of cases (95%CI 9-30%). These findings are consistent with a role for CE as a noninvasive triage investigation before DAE in patients with suspected CCD, although prospective comparative studies are required to define this sequential approach. CE may therefore be considered a first-line, noninvasive investigation in high-risk patients, with DAE reserved for targeted biopsy, evaluation of ulcerative jejunoileitis, or assessment of stricturing disease, in line with ESGE guidance [26].

Given the high DY, a central question is whether CE meaningfully influences clinical management in this population. In our study, CE findings were associated with changes in clinical management in a substantial proportion of patients across studies; however, this outcome was highly heterogeneous and should be interpreted descriptively rather than as a precise pooled estimate (Supplementary Fig. 1). The proportion of patients experiencing a management change was lower than that reported for CE in suspected general small-bowel disease, in which management changes occur in approximately 70% of patients [31]. An exploratory subgroup analysis suggested

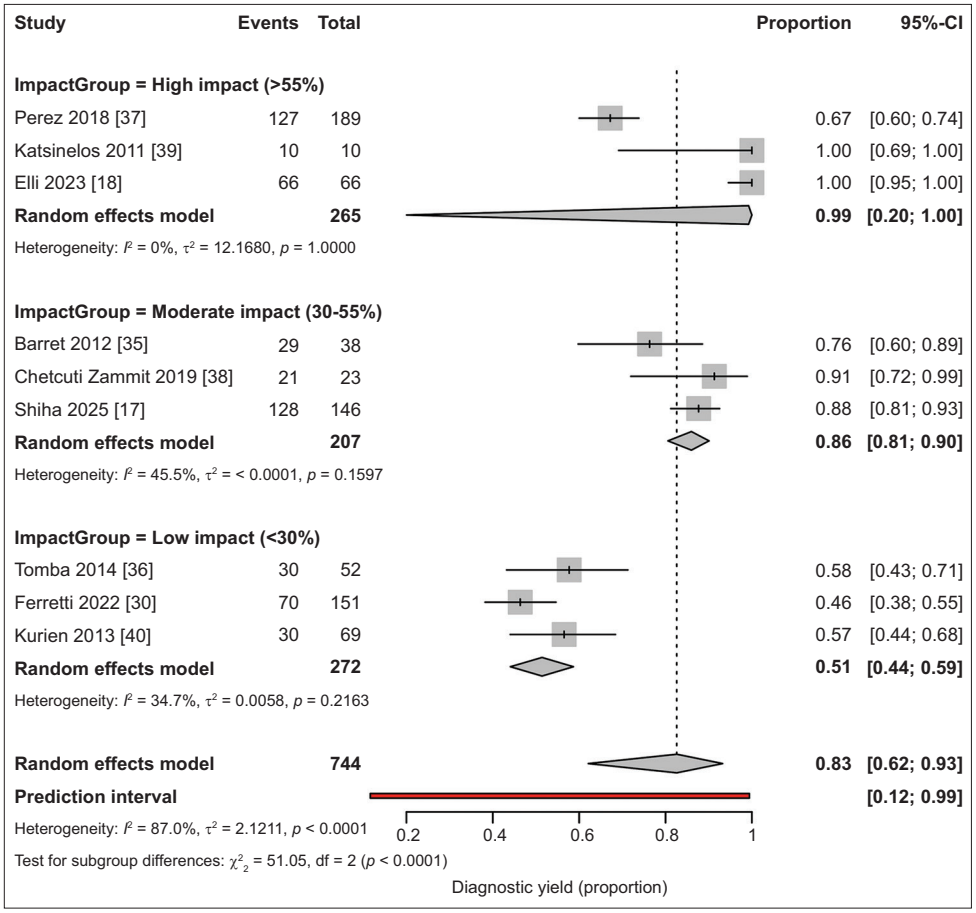



**Figure 4** Subgroup analysis of diagnostic yield stratified by reported impact of capsule endoscopy on clinical management. Studies were grouped by the proportion of patients experiencing a change in management following capsule endoscopy (high >55%, moderate 30-55%, low <30%). Diagnostic yield increased across groups, with the highest yield observed in studies reporting high clinical impact (99%, 95%CI 20-100%), compared with moderate (86%, 95%CI 81-90%) and low (51%, 95%CI 44-59%) impact groups. A statistically significant between-group difference was observed (P<0.001). However, only 9 studies reported management change and definitions varied across studies; therefore, these findings should be considered exploratory

*CI, confidence interval; $I^2$, Higgins heterogeneity statistic; $\tau^2$, between-study variance; df, degrees of freedom; $\chi^2$, chi-squared statistic*

higher DY in studies reporting greater management change. However, only 9 studies contributed data, while outcome definitions were heterogeneous, limiting the robustness of this observation. Within the subset of patients with confirmed RCD only, studies showed higher malignant DYs (6%; 95%CI 1-37%; $I^2$=0) compared with studies without RCD (2%; 95%CI 0-6%; $I^2$=0) (Supplementary Fig. 3). These findings reinforce the use of CE as a surveillance tool in this high-risk cohort [12]. However, few studies distinguished between RCD type 1 and type 2, precluding subtype-specific analyses. Although the need for surveillance in RCD is clear, the optimal interval and duration remain uncertain and represent important areas for future research. Importantly, these findings should be interpreted in the context of disease phenotype. The DY and clinical impact of CE were highest in patients with established RCD, whereas evidence supporting its use in NRCD alone remains more limited and of lower certainty.

Across a limited number of studies, patients with negative CE findings did not subsequently develop new CD-related complications during follow-up, suggesting a high negative predictive value for clinically significant small-bowel pathology. In a large multicenter European cohort, no patient with a negative examination was later found to have EATL, ulcerative jejunoileitis, or other major complications, with alternative non-celiac diagnoses identified in a substantial proportion. Therefore, a negative small-bowel CE appears clinically reassuring, with available evidence suggesting a low risk of missed celiac-related complications and frequent identification of alternative diagnoses. It may therefore support a strategy of conservative management and avoidance of further invasive investigation, although this inference is based on limited and heterogeneous data. Assessment of publication bias suggested the presence of small-study effects. Funnel plot asymmetry, supported by Egger's regression test, and subsequent trim-and-fill analysis indicated that the pooled DY may be modestly overestimated, with adjustment reducing the estimate from 82% to 79%. These findings are consistent with the recognized tendency for single-arm diagnostic studies with lower yields to

be under-reported, although the limited number of included studies constrains the reliability of formal bias assessment.

Limitations of this study include the predominance of retrospective studies, which resulted in a substantial proportion being rated as having a serious RoB, and contributed to an overall very low certainty of evidence. Missing data were frequent, with many studies omitting both negative CE examinations and information on how findings influenced subsequent management. Rare outcomes, such as malignancy, further contributed to imprecision, because multicenter studies varied in their definitions of GFD adherence and indications for CE, thereby exacerbating the already substantial heterogeneity. A potential overestimation of CCD DY cannot be excluded, as some studies lacked clarity on whether structural and malignant findings, such as strictures and malignancy, were reported as distinct entities, introducing a risk of double-counting individual patients. These methodological constraints reduce confidence in the pooled estimates and underscore the need for larger, prospective, standardized studies to better define the DY and clinical impact of CE in NRCD and RCD. Advances in artificial intelligence (AI), particularly in lesion characterization, may eventually enable the prediction of the RCD subtype directly from CE imaging, without the need for DAE and biopsy [32]. A notable methodological limitation across the included studies is the absence of a uniformly applied, validated scoring system for CE in CD, and the variation in histological reference standards. The threshold defining a "positive" examination differed substantially between study cohorts (Supplementary Table 8). This substantial heterogeneity in the definition of a positive CE examination and clinical impact constitutes a key study-level limitation, undermining direct cross-study comparability and limiting the interpretation of pooled DY estimates. Consequently, the pooled estimates should be regarded as summary measures across heterogeneous definitions of clinically significant findings, rather than precise estimates applicable to a single standardized diagnostic threshold. A further limitation is the absence of a uniformly applied reference standard across studies, with CE findings not consistently validated against histology or DAE in all participants. This limits the ability to interpret DY in relation to underlying disease status, and may have contributed to heterogeneity in reported yields. In addition, most studies were conducted in tertiary referral settings, and included highly selected populations enriched for refractory disease, raising the possibility of spectrum bias and overestimation of DY. Standardization of CE reporting criteria and reporting outcome in CD remains a priority for future prospective studies.

In conclusion, this study adds to the growing evidence on the DY of CE across NRCD, RCD and CCD cohorts, and suggests that it can meaningfully inform subsequent clinical management. Its noninvasive nature and low conversion to subsequent DAE suggest it may have a role as a second-line investigation and triage tool in patients with CxCD. However, these findings should be interpreted cautiously, given the substantial heterogeneity and predominantly very-low-certainty evidence. Future studies should evaluate the role of AI in image interpretation, although robust prospective evidence is needed before AI-assisted approaches can be integrated into routine diagnostic pathways for NRCD, RCD and CCD.

**Summary Box**

**What is already known:**

- Capsule endoscopy (CE) is recommended as a second-line investigation in suspected refractory or complicated celiac disease (CD), but no prior systematic review has evaluated its diagnostic yield specifically in this population
- Up to 30% of adults with CD have persistent symptoms, despite a gluten-free diet, and these cases carry an elevated risk of complications including ulcerative jejunitis and lymphoma

**What the new findings are:**

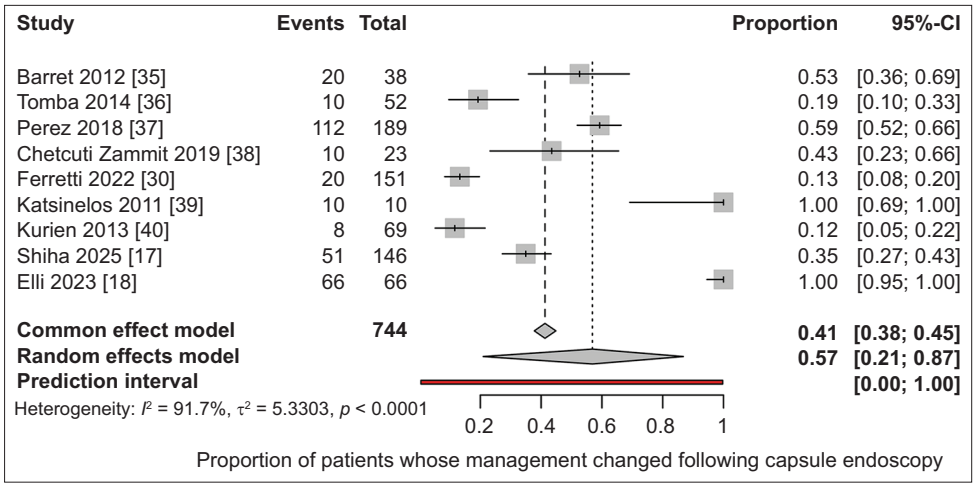
- The pooled diagnostic yield of CE in complex CD was 82% (95% confidence interval 67-91%), rising to 99% in refractory CD, with low pooled rates of ulcerative jejunitis (3%) and small-bowel malignancy (3%)
- The proportion of patients with known refractory CD and the proportion experiencing management change were significantly associated with a higher diagnostic yield, and only 18% required conversion to device-assisted enteroscopy

## References


1. Al-Toma A, Volta U, Auricchio R, et al. European Society for the Study of Coeliac Disease (ESsCD) guideline for coeliac disease and other gluten-related disorders. *United European Gastroenterol J* 2019;**7**:583-613.
2. Lebwohl B, Sanders DS, Green PHR. Coeliac disease. *Lancet* 2018;**391**:70-81.
3. Penny HA, Baggus EMR, Rej A, Snowden JA, Sanders DS. Non-responsive celiac disease: a comprehensive review from the NHS England National Centre for Refractory Celiac Disease. Nutrients 2020;**12**:216.
4. Chetcuti Zammit S, Kurien M, Sanders DS, Sidhu R. What is the role of small bowel capsule endoscopy in established celiac disease? *Clin Res Hepatol Gastroenterol* 2020;**44**:753-761.
5. Blom JJ, Gidrewicz D, Turner J, Duerksen DR, Pinto-Sánchez MI. Diagnosis and management of celiac disease. *CMAJ* 2025;**197**:E1258-E1265.
6. Malamut G, Soderquist CR, Bhagat G, Cerf-Bensussan N. Advances in nonresponsive and refractory celiac disease. *Gastroenterology* 2024;**167**:132-147.
7. Elfström P, Granath F, Ekström Smedby K, et al. Risk of lymphoproliferative malignancy in relation to small intestinal histopathology among patients with celiac disease. *J Natl Cancer Inst* 2011;**103**:436-444.
8. Malamut G, Cellier C. Refractory coeliac disease. *Curr Opin Oncol*

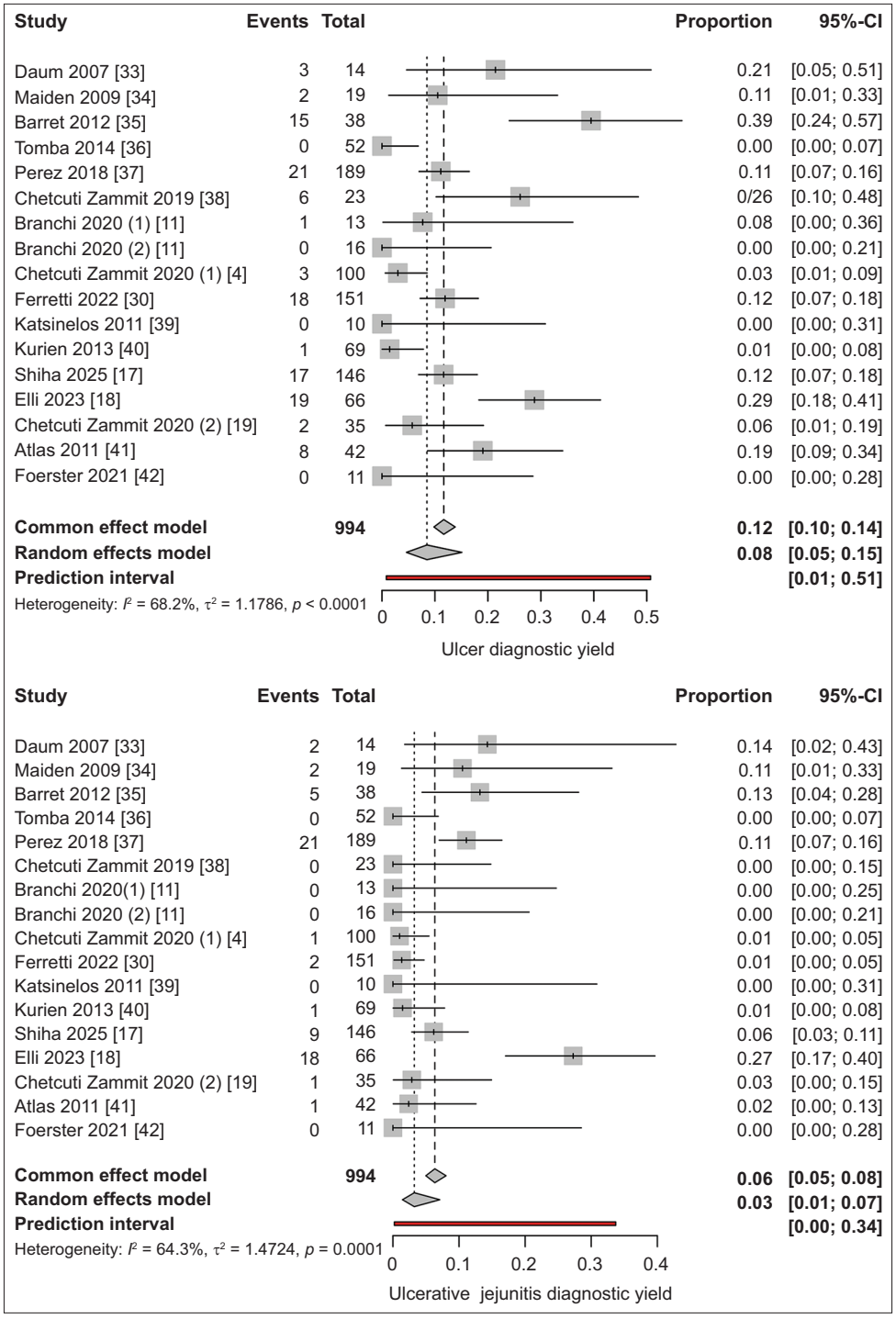
2013;**25**:445-451.
9. Nijeboer P, van Wanrooij RL, Tack GJ, Mulder CJ, Bouma G. Update on the diagnosis and management of refractory coeliac disease. *Gastroenterol Res Pract* 2013;**2013**:518483.
10. Lewis SK, Semrad CE. Capsule endoscopy and enteroscopy in celiac disease. *Gastroenterol Clin North Am* 2019;**48**:73-84.
11. Branchi F, Ferretti F, Orlando S, et al. Small-bowel capsule endoscopy in patients with celiac disease, axial versus lateral/panoramic view: Results from a prospective randomized trial. *Dig Endosc* 2020;**32**:778-784.
12. Al-Toma A, Zingone F, Branchi F, et al. European Society for the Study of Celiac Disease 2025 updated guidelines on the diagnosis and management of celiac disease in adults. Part 1: Diagnostic approach. *United European Gastroenterol J* 2025;**13**:1855-1886.
13. Elli L, Casazza G, Locatelli M, et al. Use of enteroscopy for the detection of malignant and premalignant lesions of the small bowel in complicated celiac disease: a meta-analysis. *Gastrointest Endosc* 2017;**86**:264-273.
14. Rokkas T, Niv Y. The role of video capsule endoscopy in the diagnosis of celiac disease: a meta-analysis. *Eur J Gastroenterol Hepatol* 2012;**24**:303-308.
15. Shapiro M, Niv Y. Diagnostic yield of video capsule endoscopy (VCE) in celiac disease (CD): a systematic review and meta-analysis. *J Clin Gastroenterol* 2025;**59**:598-606.
16. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *J Clin Epidemiol* 2021;**134**:178-189.
17. Shiha MG, Oka P, Nandi N, et al. Role of capsule endoscopy and double-balloon enteroscopy in the management of adult patients with coeliac disease and persisting symptoms. *Dig Liver Dis* 2025;**57**:206-212.
18. Elli L, Soru P, Roncoroni L, et al. Clinical features of type 1 and 2 refractory celiac disease: Results from a large cohort over a decade. *Dig Liver Dis* 2023;**55**:235-242.
19. Chetcuti Zammit S, Schiepatti A, Aziz I, Kurien M, Sanders DS, Sidhu R. Use of small-bowel capsule endoscopy in cases of equivocal celiac disease. *Gastrointest Endosc* 2020;**91**:1312-1321.
20. Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). *Environ Int* 2024;**186**:108602.
21. Brozek JL, Canelo-Aybar C, Akl EA, et al; GRADE Working Group. GRADE Guidelines 30: the GRADE approach to assessing the certainty of modeled evidence-An overview in the context of health decision-making. *J Clin Epidemiol* 2021;**129**:138-150.
22. Viechtbauer W. Conducting meta-analyses in R with the metafor package. In Journal of Statistical Software 2010;**36**:1-48.
23. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2024. Available from: https://www.R-project.org/ [Accessed 25 June 2026].
24. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health* 2019;**22**:153-160.
25. Wickham H, François R, Henry L, Müller K. dplyr: A Grammar of Data Manipulation. R package version 1.1.2; 2023. Available from: https://CRAN.R-project.org/package=dplyr [Accessed 25 June 2026].
26. Pennazio M, Rondonotti E, Despott EJ, et al. Small-bowel capsule endoscopy and device-assisted enteroscopy for diagnosis and treatment of small-bowel disorders: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2022. *Endoscopy* 2023;**55**:58-95.
27. Estevinho MM, Pinho R, Fernandes C, et al. Diagnostic and therapeutic yields of early capsule endoscopy and device-assisted enteroscopy in the setting of overt GI bleeding: a systematic review with meta-analysis. *Gastrointest Endosc* 2022;**95**:610-625.
28. Dionisio PM, Gurudu SR, Leighton JA, et al. Capsule endoscopy has a significantly higher diagnostic yield in patients with suspected and established small-bowel Crohn's disease: a meta-analysis. *Am J Gastroenterol* 2010;**105**:1240-1248.
29. Liao Z, Gao R, Xu C, Li ZS. Indications and detection, completion, and retention rates of small-bowel capsule endoscopy: a systematic review. *Gastrointest Endosc* 2010;**71**:280-286.
30. Ferretti F, Branchi F, Orlando S, et al. Effectiveness of capsule endoscopy and double-balloon enteroscopy in suspected complicated celiac disease. *Clin Gastroenterol Hepatol* 2022;**20**:941-949.
31. Alali AA, Alrashidi R, Allahow F, Dangi A, Alfadhli A. Predictive factors of significant findings on capsule endoscopy in patients with suspected small bowel bleeding. *Diagnostics (Basel)* 2024;**14**:2352.
32. Messmann H, Bisschops R, Antonelli G, et al. Expected value of artificial intelligence in gastrointestinal endoscopy: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. *Endoscopy* 2022;**54**:1211-1231.
33. Daum S, Wahnschaffe U, Glasenapp R, et al. Capsule endoscopy in refractory celiac disease. *Endoscopy* 2007;**39**:455-458.
34. Maiden L, Elliott T, McLaughlin SD, Ciclitira P. A blinded pilot comparison of capsule endoscopy and small bowel histology in unresponsive celiac disease. *Dig Dis Sci* 2009;**54**:1280-1283.
35. Barret M, Malamut G, Rahmi G, et al. Diagnostic yield of capsule endoscopy in refractory celiac disease. *Am J Gastroenterol* 2012;**107**:1546-1553.
36. Tomba C, Elli L, Bardella MT, et al. Enteroscopy for the early detection of small bowel tumours in at-risk celiac patients. *Dig Liver Dis* 2014;**46**:400-404.
37. Perez-Cuadrado-Robles E, Lujan-Sanchis M, Elli L, et al; Enteroscopy and Capsule Endoscopy Spanish Society Group of the Spanish Society of Digestive Endoscopy (SEED). Role of capsule endoscopy in alarm features and non-responsive celiac disease: a European multicenter study. *Dig Endosc* 2018;**30**:461-466.
38. Chetcuti Zammit S, Sanders DS, Cross SS, Sidhu R. Capsule endoscopy in the management of refractory coeliac disease. *J Gastrointestin Liver Dis* 2019;**28**:15-22.
39. Katsinelos P, Fasoulas K, Beltsis A, et al. Diagnostic yield and clinical impact of wireless capsule endoscopy in patients with chronic abdominal pain with or without diarrhea: a Greek multicenter study. *Eur J Intern Med* 2011;**22**:e63-e66.
40. Kurien M, Evans KE, Aziz I, et al. Capsule endoscopy in adult celiac disease: a potential role in equivocal cases of celiac disease? *Gastrointest Endosc* 2013;**77**:227-232.
41. Atlas DS, Rubio-Tapia A, Van Dyke CT, Lahr BD, Murray JA. Capsule endoscopy in nonresponsive celiac disease. *Gastrointest Endosc* 2011;**74**:1315-1322.
42. Foerster F, Neumann H. Panintestinal capsule endoscopy in patients with celiac disease. *Eur J Gastroenterol Hepatol* 2021;**33**:e1022-e1026.

## Supplementary material

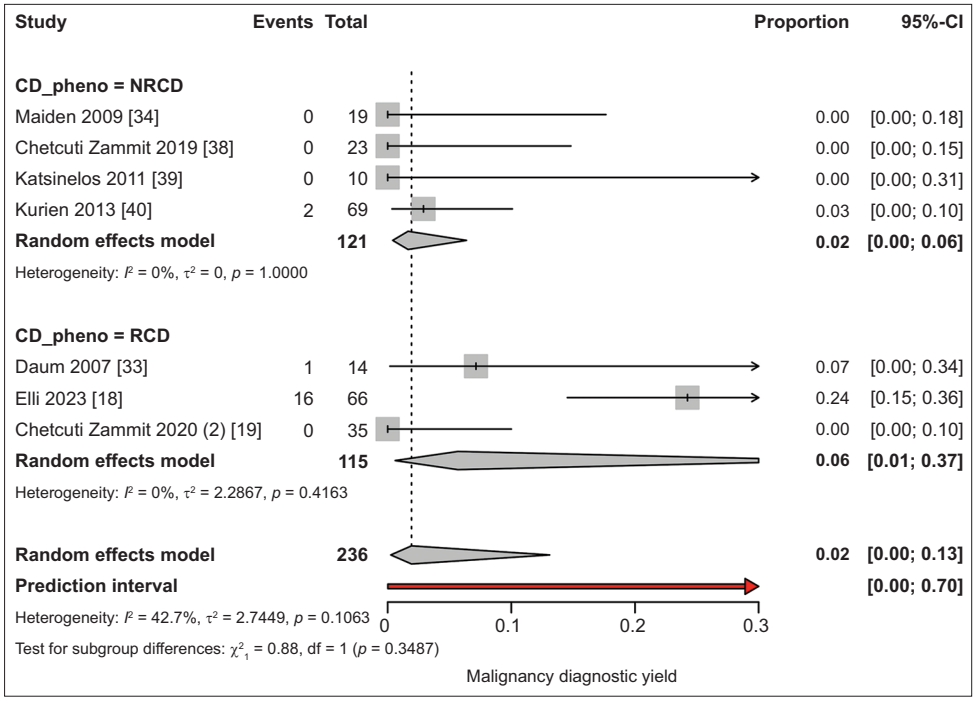



**Supplementary Figure 1** Pooled proportion of patients whose management changed following capsule endoscopy in non-responsive or refractory celiac disease

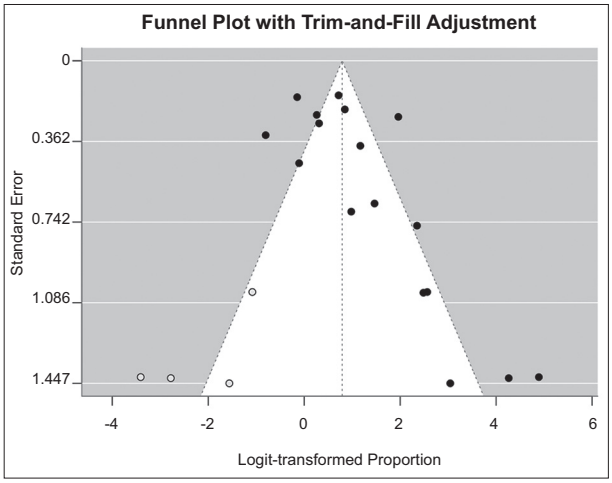
*CI, confidence interval; $I^2$, Higgins heterogeneity statistic*



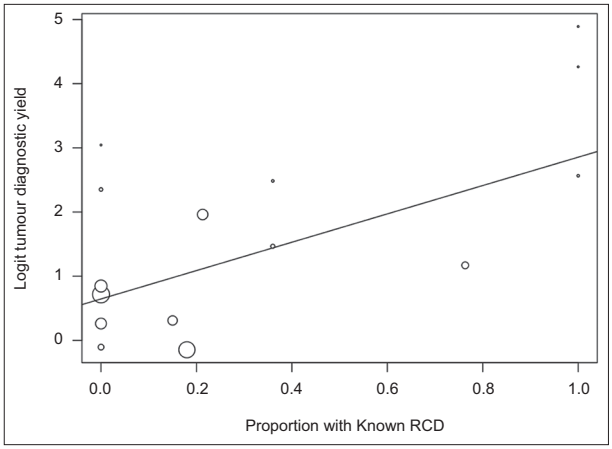

**Supplementary Figure 2** Forest plot of (A) ulcers diagnostic yield (left) and (B) ulcerative jejunitis (right)
*CI, confidence interval; $I^2$, Higgins heterogeneity statistic*



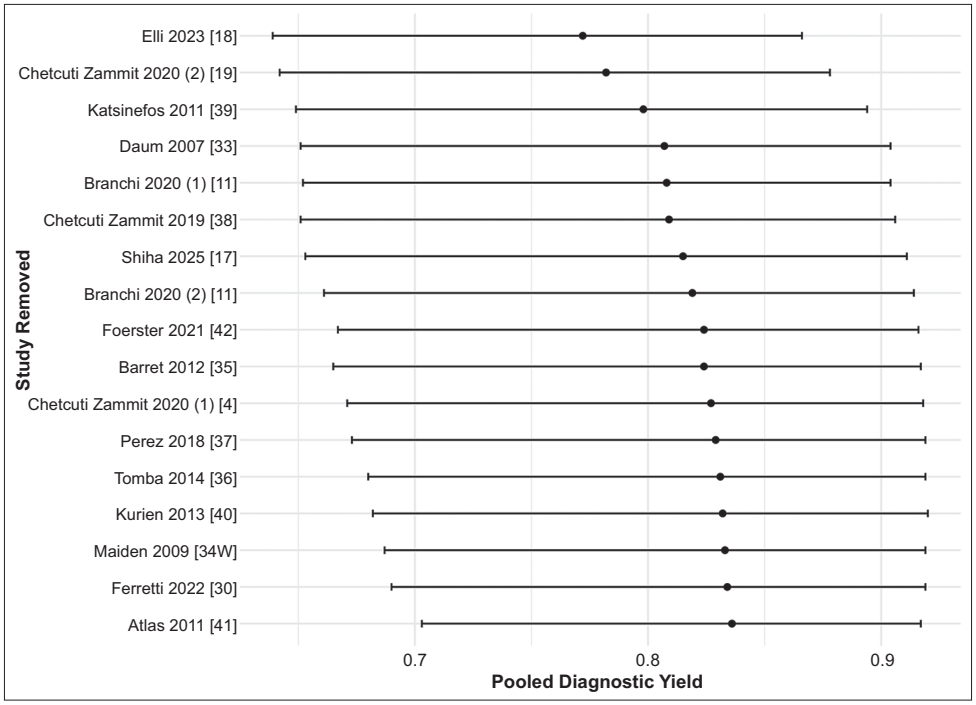

**Supplementary Figure 3** Forest plot of malignancy diagnostic yield in studies reporting RCD versus no RCD patients
*RCD, refractory celiac disease; CI, confidence interval; $I^2$, Higgins heterogeneity statistic*




**Supplementary Figure 4** The funnel plot of the selected studies reveals asymmetry. Application of the trim-and-fill method imputed 4 potentially missing studies on the left side of the funnel plot. This was generated using the "metafor" R package




**Supplementary Figure 5** Bubble plot illustrating the upward trend in tumor diagnostic yield as the proportion of participants with established refractory celiac disease (RCD) increases across studies




**Supplementary Figure 6** Leave-one-out sensitivity analysis. The graph shows the effect of excluding each study on the overall pooled effect size. Black dots represent the pooled estimates after removing 1 study, with black lines indicating the 95% confidence intervals. The gray line marks the overall pooled diagnostic yield. The minimal deviation and overlapping confidence intervals suggest that no single study significantly influenced the results

**Supplementary Table 1** PRISMA 2020 checklist

| Section and Topic | Item # | Checklist item | Location where item is reported |
|---|---|---|---|
| | | TITLE | |
| Title | 1 | Identify the report as a systematic review. | P. 1 |
| | | ABSTRACT | |
| Abstract | 2 | See the PRISMA 2020 for Abstracts checklist. | P. 2 |
| | | INTRODUCTION | |
| Rationale | 3 | Describe the rationale for the review in the context of existing knowledge. | P. 4 |
| Objectives | 4 | Provide an explicit statement of the objective (s) or question (s) the review addresses. | P. 4 |
| | | METHODS | |
| Eligibility criteria | 5 | Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses. | P. 5 |
| Information sources | 6 | Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted. | P. 5 |
| Search strategy | 7 | Present the full search strategies for all databases, registers and websites, including any filters and limits used. | Appendix |
| Selection process | 8 | Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process. | P. 5 |
| Data collection process | 9 | Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process. | P. 5 |
| Data items | 10a | List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g., for all measures, time points, analyses), and if not, the methods used to decide which results to collect. | P. 6, Appendix |
| | 10b | List and define all other variables for which data were sought (e.g., participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information. | P. 5-6 |
| Study risk of bias assessment | 11 | Specify the methods used to assess risk of bias in the included studies, including details of the tool (s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process. | P. 6 and Appendix |
| Effect measures | 12 | Specify for each outcome the effect measure (s) (e.g., risk ratio, mean difference) used in the synthesis or presentation of results. | P. 6 |
| Synthesis methods | 13a | Describe the processes used to decide which studies were eligible for each synthesis (e.g., tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)). | P. 5-6 |
| | 13b | Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions. | P. 6 |
| | 13c | Describe any methods used to tabulate or visually display results of individual studies and syntheses. | P. 6 |
| | 13d | Describe any methods used to synthesize results and provide a rationale for the choice (s). If meta-analysis was performed, describe the model (s), method (s) to identify the presence and extent of statistical heterogeneity, and software package (s) used. | P. 6 |
| | 13e | Describe any methods used to explore possible causes of heterogeneity among study results (e.g., subgroup analysis, meta-regression). | P. 6 |
| | 13f | Describe any sensitivity analyses conducted to assess robustness of the synthesized results. | P. 6, Appendix |
| Reporting bias assessment | 14 | Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases). | P. 12, Appendix |
| Certainty assessment | 15 | Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome. | P. 6, Appendix |

(*Contd...*)

**Supplementary Table 1** (*Continued*)

| Section and Topic | Item # | Checklist item | Location where item is reported |
|---|---|---|---|
| | | RESULTS | |
| Study selection | 16a | Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram. | P. 6-7, Figure 1 |
| | 16b | Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded. | P. 6-7, Figure 1 |
| Study characteristics | 17 | Cite each included study and present its characteristics. | Table 1, Appendix |
| Risk of bias in studies | 18 | Present assessments of risk of bias for each included study. | Appendix |
| Results of individual studies | 19 | For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g., confidence/credible interval), ideally using structured tables or plots. | Figure 2 and 4, Appendix |
| Results of syntheses | 20a | For each synthesis, briefly summarize the characteristics and risk of bias among contributing studies. | P. 7-9, Appendix |
| | 20b | Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g., confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect. | P. 10-12, Figures 2-3 |
| | 20c | Present results of all investigations of possible causes of heterogeneity among study results. | P. 13-14, Appendix |
| | 20d | Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results. | P. 13, Appendix |
| Reporting biases | 21 | Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed. | P. 12, Appendix |
| Certainty of evidence | 22 | Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed. | P. 7, Appendix |
| | | DISCUSSION | |
| Discussion | 23a | Provide a general interpretation of the results in the context of other evidence. | P. 14-15 |
| | 23b | Discuss any limitations of the evidence included in the review. | P. 15 |
| | 23c | Discuss any limitations of the review processes used. | P. 15 |
| | 23d | Discuss implications of the results for practice, policy, and future research. | P. 15 |
| | | OTHER INFORMATION | |
| Registration and protocol | 24a | Provide registration information for the review, including register name and registration number, or state that the review was not registered. | P. 4 |
| | 24b | Indicate where the review protocol can be accessed, or state that a protocol was not prepared. | P. 4 |
| | 24c | Describe and explain any amendments to information provided at registration or in the protocol. | N/A |
| Support | 25 | Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review. | P. 16 |
| Competing interests | 26 | Declare any competing interests of review authors. | P. 16 |
| Availability of data, code and other materials | 27 | Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review. | P. 16 |

**Supplementary Table 2** The PECO framework

| Domain | Search Terms |
|---|---|
| Population | Adult patients (age >18) with coeliac/celiac disease, including refractory and complicated forms.<br>General coeliac/celiac disease: "celiac disease", "coeliac disease", "celiac sprue", "coeliac sprue", "gluten-sensitive enteropathy", "gluten sensitive enteropathy", "gluten enteropathy"<br>Refractory/persistent disease: "refractory celiac disease", "refractory coeliac disease", "RCD", "RCD1", "RCD2", "type I/II refractory celiac/coeliac" "nonresponsive celiac/coeliac disease", "unresponsive celiac/coeliac disease", "persistent villous atrophy", "ongoing villous atrophy", "persistent enteropathy"<br>Complicated celiac disease: "complicated celiac/coeliac disease"<br>Specific complications: "ulcerative jejunitis", "ulcerative jejunoileitis", "small bowel strictures", "small intestinal strictures"<br>Malignant transformation (celiac-associated): "enteropathy-associated T-cell lymphoma", "EATL", "small bowel lymphoma" |
| Exposure | Capsule endoscopy (CE) used for diagnosis or monitoring of celiac disease: Small bowel capsule endoscopy (SBCE), Colon capsule endoscopy (CCE), Panenteric capsule endoscopy (PCE), Gastric capsule endoscopy<br>Commercial systems: PillCam, OMOM, MiroCam, EndoCapsule, NaviCam, CapsoCam |
| Comparison | Conventional diagnostic modalities used in celiac disease or its complications: Upper GI endoscopy (OGD/EGD) with duodenal biopsies, Ileocolonoscopy/colonoscopy, Push enteroscopy, balloon-assisted or device-assisted enteroscopy, CT, CT enterography, CT enteroclysis, MR enterography (MRE)/MRI, Small bowel follow-through/barium studies, Abdominal or small bowel ultrasound, Other radiological or cross-sectional imaging investigating refractory or complicated celiac disease |
| Outcome | Outcomes describing the diagnostic yield and clinical impact of CE in celiac disease, particularly refractory or complicated cases<br>Diagnostic outcomes: Diagnostic yield for lesions (ulcerative jejunitis, strictures, atrophy, neoplasia including EATL/adenocarcinoma), Diagnostic accuracy (sensitivity, specificity, predictive values)<br>Clinical impact: Management change (immunosuppression, chemotherapy, surgery, nutritional interventions), Risk stratification for complications (lymphoma, stricturing disease, ulcerative jejunitis), Symptom improvement, mucosal healing, survival, quality of life |

*PECO, Population, exposure, comparison, outcome; EATL, enteropathy-associated T-cell lymphoma; GI, gastrointestinal; OGD, esophagogastroduodenoscopy; EGD, esophagogastroduodenoscopy; CT, computed tomography; MRE, magnetic resonance enterography*

**Supplementary Table 3** Additional information for all studies

| Author [ref.] | Positive findings diagnostic yield (%) | Ulcerative jejunitis DY (%) | Malignancy DY (%) | Ulcers DY (%) | Adverse event (%) | Proportion of patients with known RCD (%) | Management change following CE (%) | Enteroscopy conversion (n) |
|---|---|---|---|---|---|---|---|---|
| Daum (2007) [33] | 92.9 | 14.3 | 7.1 | 21.4 | – | 100 | – | 2 |
| Maiden (2009) [34] | 47.4 | 10.5 | 0 | 10.5 | – | 0 | – | – |
| Barret (2012) [35] | 76.3 | 13.2 | 7.9 | 40 | – | 76.3 | 52.6 | 26 |
| Tomba (2014) [36] | 58 | 0 | 5.7 | 0 | 1.9 | 15 | 18.9 | 9 |
| Perez (2018) [37] | 67.2 | 11.1 | 3.7 | 11.1 | 0.5 | 0 | 59.3 | 25 |
| Chetcuti Zammit (2019) [38] | 91.3 | – | 0 | 26.1 | – | 0 | 42.4 | – |
| Branchi (2020) [11] | 52 | – | 0 | 4 | 0 | 36 | – | – |
| Branchi (2020) [11] | 65 | – | 0 | 0 | 0 | 36 | – | – |
| Chetcuti Zammit (2020) [4] | 70 | 1 | 1 | 3.3 | 1 | 0 | – | – |
| Ferretti (2022) [30] | 46 | 1.3 | 1.3 | 11.9 | 0.7 | 18 | 13.4 | 21 |
| Katsinelos 2011 [39] | 100 | – | 0 | 0 | 0.8 | 0 | 100 | – |
| Kurien (2013) [39] | 56.5 | 1.4 | 2.9 | 1.4 | – | 0 | 11.6 | 8 |
| Shiha (2025) [17] | 87.6 | 6.2 | 6.1 | 11.6 | 0.1 | 21.3 | 34.8 | 20 |
| Elli (2023) [18] | 100 | 27 | 24.3 | 29 | – | 100 | 100 | 25 |
| Chetcuti Zammit (2020) [19] | 51.4 | 2.9 | 0 | 5.7 | 0 | 0 | – | – |
| Atlas (2011) [41] | 31 | 2.4 | 2.4 | 19 | 0 | – | – | 1 |
| Foerster (2021) [42] | 72.7 | 0 | 0 | 0 | 0 | – | – | – |

*DY, diagnostic yield; CE, capsule endoscopy; RCD, refractory celiac disease*

**Supplementary Table 4** Clinical impact of capsule endoscopy on management in celiac disease: study-level attribution

| Domain | Specific findings/interventions | Studies reporting this |
|---|---|---|
| Pharmacological escalation | Initiation or escalation of corticosteroids (budesonide/prednisolone) following CE findings of persistent villous atrophy or inflammation<br>Use of immunosuppressive therapy (e.g., azathioprine) guided by CE findings and disease severity | (Shiha *et al*. 2025) [17] |
| | Subtype-directed therapy: corticosteroids for RCD1; immunosuppressants (azathioprine, cyclosporine, cladribine) for RCD2; chemotherapy for EATL | Ferretti *et al* (2022) [30] |
| Referral for device-assisted enteroscopy (DAE) | CE-triggered referral for tissue acquisition, confirmation of suspected lesions, or therapeutic intervention<br>CE-guided DBE confirming complications (ulceration, strictures, malignancy) and enabling therapy (e.g., dilation, capsule retrieval) | (Shiha *et al*. 2025) [17] |
| | CE-directed enteroscopy in suspected complicated CD | Ferretti *et al* (2022) [30]; Barrett *et al* (2012) [35] |
| Dietary and nutritional interventions | Reinforcement of GFD adherence following persistent villous atrophy or ongoing symptoms<br>Nutritional optimization, including supplementation and support in patients with severe disease | (Shiha *et al*. 2025) [17] |
| | Requirement for advanced nutritional support (e.g., parenteral nutrition in severe RCD) | Elli *et al* (2023) [18] |
| Surgical referral | Referral for surgery in patients with strictures or suspected malignancy not amenable to endoscopic therapy | (Shiha *et al*. 2025) [17]; Ferretti *et al* (2022) [30] |
| De-escalation/triage | Avoidance of further invasive investigations in patients with negative or low-risk CE findings | Perez-Cuadrado-Robles *et al* (2018) [37] |
| | *CE-based reclassification of disease severity leading to avoidance of unnecessary enteroscopy* | Ferretti *et al* (2022) [30] |

*CE, capsule endoscopy; RCD, refractory celiac disease; RCD1, refractory celiac disease type 1; RCD2, refractory celiac disease type 2; EATL, enteropathy-associated T-cell lymphoma; DAE, device-assisted enteroscopy; DBE, double-balloon enteroscopy; CD, celiac disease; GFD, gluten-free diet*

**Supplementary Table 5** The risk of bias assessment using the ROBINS-E tool

| Study [ref.] | Type of included studies | Bias due to confounding | Bias in participant selection | Bias in classification of exposure | Bias in deviations from exposure | Bias due to missing data | Bias in measurement of outcomes | Bias in selection of reported result | Overall RoB | Justification |
|---|---|---|---|---|---|---|---|---|---|---|
| Daum (2007) [33] | Retrospective single-center | Low | Moderate | Low | Low | Low | Low | Moderate | Low | Only 64% of examinations complete, so only 9 patients completely analyzed in the end (<10) |
| Maiden (2009) [34] | Prospective single-center | Low | Serious | Moderate | Low | Moderate | Moderate | Moderate | Serious | Single observer only. Single tertiary clinic, small sample, only those who agreed to VCE, unclear whether recruitment was consecutive, or how many eligible patients were missed |
| Barret (2012) [35] | Retrospective single-center | Low | Serious | Low | Low | Low | Moderate | Low | Serious | Selection of a non-consecutive subset of RCD patients (only 29/63) who had both CE+endoscopy within a year, these are likely more severe/complex disease |
| Tomba (2014) [36] | Prospective single-center | Low | Serious | Low | Low | Low | Low | Moderate | Serious | 10/19 patients had non-compliance with GFD+only over 50 so more likely for complications. Includes pts with ALARM Sx too (not just CD Sx) so more likely to have malignancy |
| Perez (2018) [37] | Retrospective multicenter | Low | Moderate | Low | Low | Low | Moderate | Moderate | Moderate | Only included CE patients who had 1 year follow up. Multicenter-definition of GFD adherence based on center-specific guideline. Doesn't explicitly say CE readers were blinded to patient details |
| Chetcuti Zammit (2019) [38] | Prospective single-center | Moderate | Moderate | Low | Low/ moderate | Moderate | Low | Moderate | Moderate | More about role of a second look SBCE in patients with RCD who were eligible for our study |
| Branchi (2020) [11] | Prospective single-center | Low | Moderate | Low | Low | Moderate | Low | Low | Moderate | Prospective, randomized, each patient receives both capsule types. Clear definitions and blinded expert reading. However selected, small, single-center cohort. 5 patients had missing data/incomplete examination |
| Chetcuti Zammit (2020) [4] | Prospective single-center | Moderate | Moderate | Low | Low | Moderate | Low | Moderate | Moderate | Tertiary-center, highly selected symptomatic/ RCD cohort with incomplete capsules, however robust blinded SBCE readings |

(*Contd...*)

**Supplementary Table 5** (*Continued*)

| Study [ref.] | Type of included studies | Bias due to confounding | Bias in participant selection | Bias in classification of exposure | Bias in deviations from exposure | Bias due to missing data | Bias in measurement of outcomes | Bias in selection of reported result | Overall RoB | Justification |
|---|---|---|---|---|---|---|---|---|---|---|
| Ferretti (2022) [30] | Prospective single-center | Moderate | Moderate | Low | Low | Low/ moderate | Low | Moderate | Moderate | Risk factors e.g., anemia associated with positive findings, but no adjustment done. Highly selected tertiary cohort and mixture of indications including non GFD adherence |
| Kurien (2013) [39] | Prospective single-center | Moderate | Moderate | Low | Low | Low/ moderate | Moderate | Moderate | Moderate | Tertiary single-center, videos read by 3 experienced readers, aware of clinical symptoms but blinded to serology and histology. No systematic reference standard given for findings. Completion/retention rates not reported |
| Shiha (2025) [17] | Retrospective multicenter | Low | Moderate | Low | Low | Low | Low | Moderate | Moderate | Retrospective; tertiary center, several expert capsule readers with pre-defined definitions |
| Elli (2023) [18] | Retrospective multicenter | Moderate | Moderate | Low | Low | Moderate | Moderate | Serious | Serious | Includes all RCD patients attending center, however capsule endoscopy was not performed in 6/37 patients; in five of these, CE was contraindicated owing to ▌intestinal stenosis on cross-sectional imaging (CT, MR or small-bowel ultrasound) or a positive patency test. Capsule data merged with DBE, cannot extract capsule data alone. Not recorded negative CE exams, completeness rates |
| Chetcuti Zammit (2020) [19] | Prospective single-center | Moderate | Low | Low | Low | Moderate | Low | Moderate | Low | Not all eligible patients during the period had SBCE; expert readers not blinded |
| Atlas (2011) [41] | Prospective single-center | Low | Moderate | Low | Low | Moderate | Moderate | Moderate | Moderate | Consecutive tertiary-center NRCD patients, single, unvalidated capsule reader, 24% incomplete small-bowel visualization |
| Foerster (2021) [42] | Retrospective single-center | Low | Serious | Low | Low | Moderate | Moderate | Moderate | Serious | Single-center retrospective, established CD but no mention if had CE because RCD/ NRCD etc., no indication given. Small sample. Single blinded interpreter. Complete bowel examined in only 36% |

*The assessments are categorized as: low, moderate, serious and critical risk of bias. No information should be stated if there are insufficient data.*
*ROBINS-E, Risk Of Bias In Non-randomized Studies of Exposures; RoB, risk of bias; CE, capsule endoscopy; Sx, Symptoms; SBCE, small-bowel capsule endoscopy; VCE, video capsule endoscopy; GFD, gluten-free diet; RCD, refractory celiac disease; NRCD, non-responsive celiac disease*

**Supplementary Table 6** GRADE Assessment

| Study [ref.] | Primary outcome | GRADE certainty | Justification | Secondary outcome | GRADE certainty | Justification | Overall rating |
|---|---|---|---|---|---|---|---|
| Daum (2007) [33] | DY of CE for ulcerative jejunitis/intestinal T-cell lymphoma in RCD | Very Low | Observational case series; single center; small sample size; and rare adverse events | Clinical impact, safety, feasibility, procedural success | Very low | Observational and small number; imprecision | Very low |
| Maiden (2009) [34] | DY of CE for small bowel pathology in NRCD | Very low | Observational; small sample; imprecision | Diagnostic accuracy; clinical impact; safety and technical outcomes | Very low | Same limitations; indirectness; imprecision with rare events | Very low |
| Barret (2012) [35] | DY of CE for complications of NRCD/ RCD | Very low | Observational retrospective; limited population; non-consecutive; imprecision | Diagnostic accuracy/ concordance with histology, clinical impact, safety | Very low | Same limitations; rare events | Very low |
| Tomba (2014) [36] | DY of CE detecting small-bowel malignancy in 'at risk' CD | Very low | Observational; not blinded; indirectness with population; rare events so imprecision | Clinical impact, safety, feasibility | Very Low | Observational; no comparator; imprecision with rare events of adverse effects | Very low |
| Perez (2018) [37] | DY of CE for small-bowel pathology/ complications in NRCD or alarm features | Low | Observational multicenter; population aligns | Clinical impact, safety, feasibility | Low | Observational; limited indirectness and imprecision, moderate risk of bias | Low |
| Chetcuti Zammit (2019) [38] | DY of SBCE in established CD | Very low | Observational; small sample; indirectness | Lesion/complication characterization and extent of small-bowel involvement | Very low | Observational; rare events; same risk of bias | Very low |
| Branchi (2020) [11] | DY of SBCE in suspected/known complicated/NRCD | Very low | Randomized controlled; single center, only subset of 25/508 CD patients evaluated. Mixed indications, includes non-compliant pts | Diagnostic accuracy of capsule endoscopy vs. duodenal histology for small-bowel atrophy | Very low | Observational; selection issues; readers aware of capsule type; indirectness and imprecision | Very low |
| Chetcuti Zammit (2020) [4] | DY in NRCD/RCD/ complicated CD | Very low | Observational; bias with pt selection, indirectness and imprecision, rare neoplastic events | Diagnostic accuracy (sensitivity/ specificity vs duodenal histology); completion rate; retention/adverse events | Low | Observational with rare events, so imprecision; risk of bias with selected cohort. | Very low |
| Ferretti (2022) [30] | DY in NRCD/RCD/ complicated CD | Low | Observational; mixed indications, highly selective tertiary population; no comparator; modest sample size | Diagnostic accuracy, other secondary outcomes e.g., lesion/complication characterization, clinical impact, completion, retention/ adverse events | Very low | Same risk of bias; rare adverse events so imprecision and indirectness. | Very low |

(*Contd...*)

**Supplementary Table 6** (*Continued*)

| Study [ref.] | Primary outcome | GRADE certainty | Justification | Secondary outcome | GRADE certainty | Justification | Overall rating |
|---|---|---|---|---|---|---|---|
| Kurien (2013) [39] | DY of CE for significant small-bowel pathology in non-responsive CD | Very low | Observational with rare adverse events in a single center so imprecision. | Characterization of CD-related complications. Clinical impact/change in management | Very low | Same risk of bias, rare events so imprecision. Sparse and mostly descriptive data on management change. No comparator, no data on completion | Very low |
| Shiha (2025) [17] | DY for small bowel complications in NRCD/RCD/complicated CD | Very low | Observational, retrospective single-center; rare events; imprecision with small number events; mixed CD population including seronegative CD/RCD/NRCD | Completion rate, bowel prep, adverse events | Very low | Risk of bias, indirectness and imprecision | Very low |
| Elli (2023) [18] | DY for small bowel lesions/complications in RCD | Very low | Retrospective; some pts not examined; no specific stats for capsule; indirectness, imprecision | Lesion/complication characterization | Very low | Same issues; does not separate capsule from DBE | Very low |
| Chetcuti Zammit (2020) [19] | DY in equivocal/seronegative CD | Very low | Observational; unblinded outcome assessment; indirectness with population | Diagnostic accuracy; clinical impact; completion rate; complications | Very low | Rare events; no comparator | Very low |
| Atlas (2011) [41] | DY in NRCD | Very low | Observational; risk of bias; imprecision | Specific complications found | Very low | Same limitations, rare events so imprecision | Very low |
| Foerster (2021) [42] | DY of capsule in CD | Very low | Observational, risk of bias, serious impreciseness, indirectness with population selection | Completion rate, capsule retention, adverse events, clinical impact | Very low | Same limitations, rare events with complications and clinical impact | Very low |

The assessments are categorized as: very low, low, moderate, high. No information should be stated if there is insufficient data

*GRADE, Grading of Recommendations, Assessment, Development and Evaluations; DBE, Double-balloon Enteroscopy; DY, diagnostic yield; CE, capsule endoscopy; RCD, refractory celiac disease; NRCD, non-responsive celiac disease; CD, celiac disease*

**Supplementary Table 7** Meta-regression results

| Covariates | Regression | Beta | R2† | SE | P-value |
|---|---|---|---|---|---|
| Mean age | Univariate | −0.0880 | 0.000 | 0.089 | 0.320 |
| Male sex (%) | Univariate | 3.169 | 0.000 | 3.079 | 0.303 |
| Median duration of disease | Univariate | −0.2293 | 0.000 | 0.318 | 0.471 |
| Ulcers DY | Univariate | 1.840 | 0.000 | 2.698 | 0.495 |
| Ulcerative jejunitis DY | Univariate | 6.661 | 0.017 | 4.763 | 0.162 |
| Tumor DY | Univariate | 11.322 | 0.134 | 6.417 | 0.078 |
| Completion rate | Univariate | 0.749 | 0.092 | 2.632 | 0.776 |
| Bowel adequacy | Univariate | −1.5702 | 0.639 | 1.781 | 0.378 |
| Impacting subsequent management | Univariate | 3.462 | 0.523 | 1.195 | 0.004* |
| Patients with persistent elevated tTG | Univariate | −0.7953 | 0.000 | 1.411 | 0.573 |
| Patients with Known RCD | Univariate | 2.210 | 0.436 | 0.799 | 0.006* |
| Patients with persistent symptoms | Univariate | −0.9688 | 0.087 | 0.973 | 0.320 |
| Covariates | Regression | Estimate | SE | Z-Value | P-value |
| Mean age | Multivariate | 0.372 | 0.274 | 1.3573 | 0.1747 |
| Ulcers DY | Multivariate | 6.6115 | 4.851 | 1.3629 | 0.1729 |
| Ulcerative jejunitis DY | Multivariate | -37.2288 | 42.947 | -0.8669 | 0.386 |
| Tumor DY | Multivariate | -0.3513 | 21.8788 | -0.0161 | 0.9872 |
| Male sex | Multivariate | -3.2957 | 4.306 | -0.7654 | 0.444 |
| Capsule M2A capsule | Multivariate | 2.1192 | 3.8318 | 0.5531 | 0.5802 |
| Capsule mixed capsules | Multivariate | 7.3895 | 5.5212 | 1.3384 | 0.1808 |
| Capsule not specified | Multivariate | 7.0843 | 4.0238 | 1.7606 | 0.0783 |
| Capsule Pillcam SB | Multivariate | 1.0183 | 1.4871 | 0.6847 | 0.4935 |
| Known RCD | Multivariate | 6.2476 | 2.508 | 2.491 | 0.0127* |
| Symptoms | Multivariate | 0.0904 | 2.5987 | 0.0348 | 0.9722 |

†$R^2$ denotes the proportion of between-study variance ($\tau^2$) explained by the covariate relative to the intercept-only model; negative values were truncated to 0
*DY, diagnostic yield (proportion of patients with the finding)*
Intercept not shown. *$P<0.05$
*DY, diagnostic yield; SE, standard error; RCD, refractory celiac disease; tTG, tissue transglutaminase antibody*

**Supplementary Table 8** Heterogeneity in the operationalization of a positive capsule endoscopy examination across included studies

| Study [ref.] | Definition of a positive capsule endoscopy examination |
|---|---|
| Chetcuti Zammit *et al* [38] | Presence and extent of macroscopic features (e.g., villous atrophy, scalloping, fissuring, mosaic pattern), with emphasis on quantifying the proportion of affected small bowel mucosa |
| Ferretti *et al* [30] | Findings categorized into celiac disease-related abnormalities and complications (including ulceration and mass lesions), with recognition of non-specific findings |
| Perez-Cuadrado-Robles *et al* [37] | Findings interpreted in relation to clinical context and disease severity |
| Branchi *et al* [11] | Image interpretation performed using predefined criteria and/or expert consensus among readers |

## Search strings used in each database:

### Embase Classic + Embase

Search Strategy:

1. exp Capsule Endoscopy/OR capsule endoscop* OR video capsule endoscop* OR video capsule colonoscop* OR wireless capsule endoscop* OR wireless capsule colonoscop* OR capsule colonoscop* OR pillcam*
2. exp Colonoscopy/OR colonoscop* OR exp Sigmoidoscopy/ OR sigmoidoscop* OR endoscop* OR exp Endoscopy/OR exp Endoscopy, Gastrointestinal/OR Colonography, Computed Tomographic/OR colonograph* OR exp Tomography, X-Ray Computed/OR (CT scan OR CT colonoscopy* OR computer tomography colonoscop* OR computed tomography colonoscop* OR computed tomography scan) OR virtual colonoscop* OR reinvestigat* OR conver* OR (subsequent AND (investigation* OR endoscopy* OR colonoscop*))
3. 1 AND 2
4. Limit to publication years 2000–2025
5. Limit to publication type "Article"
6. Diagnostic accuracy OR clinical article OR clinical trial OR comparative study

*Final Records: 2,125*

### Ovid MEDLINE(R) ALL

Search Strategy:

1. exp Capsule Endoscopy/OR capsule endoscop* OR video capsule endoscop* OR video capsule colonoscop* OR wireless capsule endoscop* OR wireless capsule colonoscop* OR capsule colonoscop* OR pillcam*
2. exp Colonoscopy/OR colonoscop* OR exp Sigmoidoscopy/ OR sigmoidoscop* OR endoscop* OR exp Endoscopy/OR exp Endoscopy, Gastrointestinal/OR Colonography, Computed Tomographic/OR colonograph* OR exp Tomography, X-Ray Computed/OR (CT scan OR CT colonoscopy* OR computer tomography colonoscop* OR computed tomography colonoscop* OR computed tomography scan) OR virtual colonoscop* OR reinvestigat* OR conver* OR (subsequent AND (investigation* OR endoscopy* OR colonoscop*))
3. 1 AND 2
4. Limit to publication years 2000–2025
5. Limit to publication type "Clinical Trial", "Observational Study", "Comparative Study", "Evaluation Study", "Meta-Analysis", "Validation Study", "Controlled Clinical Trial", "Randomized Controlled Trial"

*Final Records: 950*

### Pubmed (Search Strategy example):

(Celiac Disease[MeSH Terms] OR Sprue, Celiac[MeSH Terms]OR Enteropathy, Gluten-Sensitive[MeSH Terms] OR celiac[tiab] OR coeliac[tiab] OR ((celiac[tiab] OR coeliac[tiab]) AND sprue[tiab]) OR (gluten[tiab] AND enteropathy[tiab]) OR "refractory celiac disease"[tiab] OR "refractory coeliac disease"[tiab] OR (refractory[tiab] AND (celiac[tiab] OR coeliac[tiab])) OR "nonresponsive celiac disease"[tiab] OR "non-responsive celiac disease"[tiab] OR "nonresponsive coeliac disease"[tiab] OR "non-responsive coeliac disease"[tiab] OR "unresponsive celiac disease"[tiab] OR "unresponsive coeliac disease"[tiab] OR "persistent enteropathy"[tiab] OR (villous[tiab] AND atrophy[tiab] AND (persistent[tiab] OR ongoing[tiab])) OR "complicated celiac disease"[tiab] OR "complicated coeliac disease"[tiab] OR "ulcerative jejunitis"[tiab] OR "ulcerative jejunoileitis"[tiab] OR (("small bowel"[tiab] OR "small intestine"[tiab] OR "small intestinal"[tiab]) AND (stricture[tiab] OR strictures[tiab])) OR "enteropathy-associated T-cell lymphoma"[tiab] OR "enteropathy associated T cell lymphoma"[tiab] OR EATL[tiab] OR ("small bowel lymphoma"[tiab] AND (celiac[tiab] OR coeliac[tiab] OR enteropathy[tiab])))

AND

("Adult"[MeSH Terms] OR adult[tiab] OR adults[tiab])

AND

("Capsule Endoscopy"[MeSH] OR "capsule endoscopy"[tiab] OR "wireless capsule endoscopy"[tiab] OR SBCE[tiab] OR CCE[tiab] OR PCE[tiab] OR "small bowel capsule"[tiab] OR "colon capsule"[tiab] OR "panenteric capsule"[tiab] OR "gastric capsule"[tiab] OR "capsule imaging"[tiab] OR "capsule colonoscopy"[tiab]OR PillCam[tiab] OR OMOM[tiab] OR MiroCam[tiab] OR EndoCapsule[tiab] OR NaviCam[tiab] OR CapsoCam[tiab])

AND

("Colonoscopy"[MeSH] OR "Endoscopy, Gastrointestinal"[MeSH] OR "Magnetic Resonance Imaging"[MeSH] OR "Tomography, X-Ray Computed"[MeSH] OR "Ultrasonography"[MeSH] OR "Radiology"[MeSH] OR "Fluoroscopy"[MeSH] OR colonoscopy[tiab] OR ileocolonoscopy[tiab] OR gastroscopy[tiab] OR EGD[tiab] OR OGD[tiab] OR "upper endoscopy"[tiab] OR endoscopy[tiab] OR enteroscopy[tiab] OR "push enteroscopy"[tiab] OR "balloon-assisted enteroscopy"[tiab] OR "device-assisted enteroscopy"[tiab] OR "double-balloon enteroscopy"[tiab] OR "single-balloon enteroscopy"[tiab] OR MRI[tiab] OR "magnetic resonance imaging"[tiab] OR "MR enterography"[tiab] OR CT[tiab] OR "computed tomography"[tiab] OR "CT enterography"[tiab] OR "CT colonography"[tiab] OR "CT colonoscopy"[tiab] OR ultrasound[tiab] OR US[tiab] OR "abdominal imaging"[tiab] OR "radiologic imaging"[tiab] OR "radiological imaging"[tiab] OR fluoroscopy[tiab] OR "small bowel follow-through"[tiab] OR enteroclysis[tiab] OR "barium study"[tiab] OR "contrast imaging"[tiab])

AND

("Sensitivity and Specificity"[MeSH] OR "Treatment Outcome"[MeSH] OR "Clinical Decision-Making"[MeSH] OR "diagnostic yield"[tiab] OR "diagnostic accuracy"[tiab] OR sensitivity[tiab] OR specificity[tiab] OR "predictive value"[tiab] OR "diagnostic performance"[tiab] OR "change in management"[tiab] OR "treatment decision"[tiab] OR "clinical outcome"[tiab] OR "symptom resolution"[tiab] OR "therapeutic impact"[tiab] OR "clinical utility"[tiab] OR "quality of life"[tiab])