Gastroenterology 2023;165:1367–1399

GUIDELINES

AGA Clinical Practice Guideline on the Role of Biomarkers for the Management of Crohn's Disease

Ashwin N. Ananthakrishnan,[1,*] **Jeremy Adler**,[2,3,*] **Karen A. Chachu**,[4,*] Nghia H. Nguyen,[5] Shazia M. Siddique,[6,7] Jennifer M. Weiss,[8] Shahnaz Sultan,[9] Fernando S. Velayos,[10] Benjamin L. Cohen,[11] and Siddharth Singh,[12,13] on behalf of the AGA Clinical Guidelines Committee*

[1]Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts; [2]Division of Pediatric Gastroenterology, C.S. Mott Children's Hospital, Michigan Medicine, University of Michigan, Ann Arbor, Michigan; [3]Susan B. Meister Child Health Evaluation and Research Center, University of Michigan, Ann Arbor, Michigan; [4]Division of Gastroenterology, Department of Medicine, Duke University, Durham, North Carolina; [5]Division of Gastroenterology, Kaiser Permanente Medical Group, Riverside, California; [6]Division of Gastroenterology, Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; [7]Center for Evidence-Based Practice, University of Pennsylvania Health System, Philadelphia, Pennsylvania; [8]Division of Gastroenterology and Hepatology, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin; [9]Division of Gastroenterology, Hepatology, and Nutrition, University of Minnesota, Minneapolis, Minnesota, and Veterans Affairs Healthcare System, Minneapolis, Minnesota; [10]Division of Gastroenterology, Kaiser Permanente Medical Group, San Francisco, California; [11]Division of Gastroenterology, Hepatology, and Nutrition, Digestive Disease Institute, Cleveland Clinic, Cleveland, Ohio; [12]Division of Gastroenterology and Hepatology, University of California San Diego, La Jolla, California; and [13]Division of Biomedical Informatics, Department of Medicine, University of California San Diego, La Jolla, California

BACKGROUND & AIMS: Biomarkers are used frequently for evaluation and monitoring of patients with Crohn's disease (CD). This American Gastroenterological Association (AGA) guideline is intended to support practitioners in decisions about the use of biomarkers for the management of CD. **METHODS:** A multidisciplinary panel of content experts and guideline methodologists used the Grading of Recommendations Assessment, Development and Evaluation framework to formulate patient-centered clinical questions and review evidence on the performance of fecal calprotectin, serum C-reactive protein (CRP), and Endoscopic Healing Index in patients with established CD who were asymptomatic, had symptoms of varying severity, or were in surgically induced remission. Biomarker performance was assessed against the gold standard of endoscopic activity, defined as a Simple Endoscopic Score for Crohn's Disease $\geq$3. The panel used the Grading of Recommendations Assessment, Development and Evaluation Evidence-to-Decision framework to develop recommendations for use of biomarkers in various settings. Implementation considerations were formulated for each recommendation to inform clinical practice. **RESULTS:** The guideline panel made 11 conditional recommendations. In patients with CD in symptomatic remission, the panel suggests use of a biomarker- and symptom-based monitoring strategy over symptoms alone. In patients in symptomatic remission, a fecal calprotectin $<$150 μg/g and normal CRP rules out active inflammation, avoiding endoscopic evaluation for assessment of disease activity. However, elevated biomarkers in this setting merit confirmation with endoscopy before treatment adjustment. In patients with CD with mild symptoms, neither normal nor elevated biomarkers alone are sufficiently accurate to determine endoscopic activity. In patients with CD with moderate to severe symptoms, elevated fecal calprotectin or serum CRP suggests endoscopic activity, precluding routine endoscopic assessment for disease activity. In patients with CD in surgically induced remission in low-risk patients on pharmacologic prophylaxis, a normal fecal calprotectin reliably rules out endoscopic recurrence. In other postoperative settings, the panel suggests endoscopic assessment for establishing postoperative recurrence. **CONCLUSIONS:** In patients with CD, fecal calprotectin and serum CRP can inform disease management in both asymptomatic and symptomatic disease. Discordance between symptom assessment and biomarker value may merit endoscopic evaluation for confirmation of status of disease activity.

Keywords: Inflammatory Bowel Disease; Monitoring; Endoscopic Remission; Treat to Target; Evidence Synthesis.

Inflammatory bowel diseases (IBDs), comprising Crohn's disease (CD) and ulcerative colitis (UC), are rising in incidence and prevalence worldwide.[1,2] They often have an onset in young adulthood and are characterized by a protracted relapsing–remitting course with progressive permanent bowel damage.[3] The therapeutic armamentarium for CD has expanded over the past decade with multiple new mechanisms of action. Despite such progress,

*Authors share co-first authorship.

Abbreviations used in this paper: **AGA, American Gastroenterological Association; AP, abdominal pain; CD, Crohn's disease; CDEIS, Crohn's Disease Endoscopic Index of Severity; CRP, C-reactive protein; EHI, Endoscopic Healing Index; FN, false negative; FP, false positive; GRADE, Grading of Recommendations Assessment, Development and Evaluation; IBD, inflammatory bowel disease; PRO2, 2-item patient-reported outcomes; PRO3, 3-item patient-reported outcomes; RCT, randomized controlled trial; SES-CD, Simple Endoscopic Score for Crohn's Disease; TN, true negative; TP, true positive; UC, ulcerative colitis.**

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2023.09.029

nearly one-half of individuals with CD require at least 1 surgery over the course of their illness.[4,5] To durably modify the natural history of CD, an important concept that has emerged is the need for early institution of effective treatment followed by confirmation of attainment of therapeutic target in order to improve long-term outcomes and prevent disease-related disability.[6–9] Cross-sectionally, symptoms in CD have correlated poorly with endoscopic disease activity, making symptom-based disease activity assessment suboptimal.[10–12] Thus, objective assessment of inflammation has traditionally relied on endoscopic assessment of disease activity and demonstration of mucosal healing. In a population-based cohort from Norway, patients with CD who achieved endoscopic healing at 1 year had superior long-term outcomes up to 5 years after diagnosis, with a reduced need for corticosteroids and fewer CD-related hospitalizations.[13] However, reliance solely on endoscopy for repeated assessment of disease activity is limited by cost and resource utilization, invasiveness, and reduced patient acceptability. In the CALM (Effect of Tight Control Management on Crohn's Disease) trial comparing a symptom-based therapeutic strategy with a biomarker-based strategy, the use of frequent biomarker measurement to guide therapy escalation was associated with improved patient outcomes over 2 years.[14] The performance of serum and fecal biomarkers of disease activity, as well as robust determination of thresholds that can function as surrogates of endoscopic activity assessment, have not been examined comprehensively, leading to significant variability in clinical practice in optimal use of these biomarkers.

Objective

The objective of this guideline was to inform the role of commonly used serum and fecal biomarkers as surrogates for endoscopic disease activity for both cross-sectional assessment and longitudinal monitoring of patients with an established diagnosis of CD. The scope of this guideline was restricted to biomarkers that are widely available commercially within the United States. This guideline also separately examined the predictive value of biomarkers for assessment of postoperative recurrence in CD. The panel did not examine the role of biomarkers in the diagnostic pathway for patients with suspected CD. The role of biomarkers in UC were also examined in a recent guideline.[15]

Target Audience

The target audience for these guidelines includes gastroenterology health care professionals; primary care, emergency, and urgent care providers; patients; and policy makers. Recommendations are provided for common clinical scenarios in typical patients with CD. However, individual patients may have unique circumstances that must be accounted for when implementing these guidelines. Each recommendation in this guideline is accompanied by key implementation considerations and qualifying remarks that should be considered an integral part of the recommendation statement and should not be omitted. Discussions about benefits and harms are important in shared decision making, particularly for conditional recommendations when patient values and specific tradeoffs are important to consider.

Methods

Overview

This document represents the official recommendations of the American Gastroenterological Association (AGA) and was developed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) framework for diagnostic tests and strategies and adheres to best practices in guideline development, as outlined by the National Academy of Medicine (formerly Institute of Medicine).[16] The development of this guideline was fully funded by the AGA Institute.

Guideline Panel Composition and Conflict of Interest

Members of the guideline panel and evidence synthesis panel were selected on the basis of their clinical and methodological expertise after a thorough vetting process. Panel members disclosed all potential conflicts of interest. Conflicts were managed according to AGA policies and National Academy of Medicine and Guidelines International Network standards. Guideline chair (K.A.C.) and co-chair/senior methodologist (S. Singh) had no direct conflicts of interest. No panel members were excluded due to a disqualifying conflict. The evidence synthesis panel consisted of 2 content experts with expertise in CD (A.N.A., J.A.), a senior guideline methodologist with expertise in evidence synthesis and GRADE (S. Singh), and 2 junior guideline methodologists (N.H.N, S.M.S.). The guideline panel was multidisciplinary, consisting of a general gastroenterologist (J.M.W.), gastroenterologists with expertise in CD (K.A.C., B.L.C., F.S.V.), and guideline methodologists (S. Singh, S. Sultan, N.H.N, S.M.S.). The input of a patient representative on the role of biomarkers in the management of IBD was also considered in framing recommendations. A full list of conflicts can be accessed at AGA's National Office in Bethesda, MD.

Scope

The guideline panel defined biomarkers as molecules that are quantifiable in tissue, blood, stool, or urine and represent an underlying biological disease process.[17] Biomarkers have been studied in CD in various clinical contexts, including assessing likelihood of a diagnosis of CD in patients with suggestive symptoms; predicting clinical course of CD, including need for surgery; development of stricturing or penetrating disease; and quantifying disease activity. The panel focused on biomarkers that are widely used for assessing disease activity and making treatment decisions, measurable in easily accessible tissue or body fluid compartments, and commercially available in the United States. We examined the performance of individual biomarkers in unselected cohorts of patients with CD, as well as for initial assessment of postoperative recurrence of CD after surgically induced remission. The panel examined the cross-sectional performance of each biomarker against endoscopic assessment of disease activity as the gold standard. Biomarkers with demonstrated utility in research studies only, but not available for widespread commercial use, were outside

the scope of this guideline. We also did not examine biomarkers that have been developed solely for prediction of likelihood of disease progression, such as development of stricturing or penetrating disease. However, the panel examined performance of biomarkers that may predict future disease activity in the context of longitudinal monitoring. Based on these criteria, we focused on serum C-reactive protein (CRP), fecal calprotectin, and the Endoscopic Healing Index (EHI, Monitr).

Formulation of Clinical Questions

Through an iterative process, the guideline and evidence synthesis panels developed focused clinical questions deemed relevant for clinical practice to be addressed in the guideline (Table 1). These related to diagnostic accuracy and utility of commonly used serum or stool biomarkers that are commercially available. Each guideline statement was derived from a focused clinical question that comprised a well-defined statement using the PICO (patients, intervention, comparator, and outcome) format. These statements were used to formulate the study inclusion and exclusion criteria for review, guided the literature search, and informed the final guideline recommendations. The AGA Governing Board approved the final set of questions in October 2021.

Search Strategy

An experienced medical librarian conducted a comprehensive search of the following databases (Ovid Medline In-Process & Other Non-Indexed Citations, Ovid MEDLINE, Embase, and Wiley Cochrane Library) from inception to November 21, 2021, using a combination of controlled vocabulary terms supplemented with keywords (Supplementary Table 1); an updated search of Ovid MEDLINE was performed on September 1, 2022. The search was limited to English language and humans. The bibliographies of prior guidelines and the included references were searched to identify relevant studies that may have been missed. In addition, content experts helped identify any additional or ongoing studies.

Study Selection, Data Abstraction, and Statistical Analysis

We included randomized controlled trials (RCTs) or observational studies of diagnostic accuracy that met the following inclusion criteria: performed in patients with CD; provided adequate description of biomarker (ie, CRP, fecal calprotectin, and EHI), with cutoff corresponding to detection of endoscopically active CD (generally corresponding to a Simple Endoscopic Score for Crohn's Disease [SES-CD] score >3 [mild to severe inflammation]); with ileocolonoscopy as gold standard (or magnetic resonance enterography, video capsule endoscopy, or balloon-assisted enteroscopy for patients with small bowel CD not adequately examined on colonoscopy); and provided sufficient data to allow estimation of diagnostic accuracy of the biomarker for detection of endoscopic activity. For wider applicability and generalizability, we preferentially chose cutoffs most commonly used in clinical practice. These cutoffs were as follows: CRP: 5 ± 5 mg/L or 0.5 ± 0.5 mg/dL; fecal calprotectin: 250 ± 50 μg/g, 150 ± 50 μg/g and 50 ± 50 μg/g; and EHI: > 50 or < 20.

We abstracted data on patient population and phenotype, biomarker, reference standard outcome, and test performance for each eligible study. Paired values of sensitivity and specificity were pooled using a bivariate regression random-effects model proposed by Reitsma et al[18] using STATA, version 14.0 software (StataCorp, College Station, TX). Statistical assessment of heterogeneity was performed using the inconsistency index (I^2), which estimates what proportion of total variation across studies was due to heterogeneity rather than chance.[19]

Outcomes of Interest and Illustrative Clinical Scenarios

For PICOs focusing on biomarker cutoffs to either detect or rule out mild to severe endoscopic activity, the preferred outcome was direct consequences on patient-important outcomes (ie, implications of true positive [TP], false positive [FP], true negative [TN], false negative [FN] results for patients, see below). As none of the studies assessed these outcomes directly, we used TP, FP, TN, and FN rates as surrogate outcomes and inferred downstream consequences on patient-important outcomes. We opted to focus on detection of mild to severe endoscopic activity (SES-CD > 3) rather than detection of moderate to severe endoscopic activity only because most studies reported the performance of biomarkers at this cutoff and treatment adjustments in CD are recommended in response to presence of any ulcers, rather than focusing only on patients with SES-CD >6.[20–22]

For questions focused on ruling out endoscopically active CD, our outcome was minimizing rates of FN (ie, patients incorrectly labeled as being in remission when they actually have active endoscopic inflammation) to a level <5% in general, with reasonable rates of TP, FP, and TN (Supplementary Figure 1). For questions focused on detecting endoscopic activity, our outcome was minimizing rates of FP (ie, patients incorrectly labeled as having active endoscopic inflammation when their disease is actually in remission) (Supplementary Figure 2). The threshold of 5% FN and FP rates is similar to that used in the UC guideline and was consistent with patient preference for choosing stool-based biomarkers over endoscopic assessment for monitoring inflammation.[23]

Overall TP, FP, TN, and FN rates are dependent on pretest probability. We derived illustrative prevalence of any endoscopic activity (SES-CD >3) based on a combination of abdominal pain (AP) and stool frequency score, 2 of the most commonly used patient-reported outcomes, which were used to calculate 2-item patient-reported outcomes (PRO2) disease activity scores. Prevalence of any endoscopic activity (SES-CD ≥3) and of endoscopic remission (SES-CD <3) for different combinations of cutoffs of PRO2 (based on AP and stool frequency score) at varying time points after treatment initiation or adjustment were derived from existing literature based on individual participant data from phase 2 and 3 clinical trial programs of biologic agents in patients with moderate to severely active CD (unpublished data), as well as referral center observational cohorts with prospective assessment of clinical disease scores and endoscopic activity.[24]

For our analysis, we used 4 illustrative scenarios, 2 pertaining to patients with CD in symptomatic remission and 2 for those with active symptoms:

- Low pretest probability of having endoscopically active inflammation comprised asymptomatic patients with CD (PRO2 <8, 3 or fewer very loose/watery stools per day, and absent or mild AP; or 3-item PRO [PRO3] <13). This population was further subdivided into 2 subgroups based on whether the patient had recently (within 3 years)

Table 1. Focused Questions and Corresponding PICO (Patients, Intervention, Comparator, and Outcome) Questions Addressed in These Guidelines

Variable	Focused question	Patients	Intervention (threshold?)	Comparator	Outcome
Patients with CD in symptomatic remission					
Question 1	In patients with CD in symptomatic remission, is interval biomarker-based monitoring superior to symptom-based monitoring to improve long-term outcomes?	Patients with established CD in symptomatic remission	Interval biomarker-based monitoring	Interval symptom-based monitoring	Maintaining clinical remission at 12 mo and beyond
Question 2[a]	In patients with CD in symptomatic remission, at what fecal calprotectin, serum CRP, and EHI cutoff can we accurately rule out active inflammation, obviating routine endoscopic assessment?	Patients with established CD in symptomatic remission, with recent confirmation of endoscopic remission (within 3 y prior, without change in therapy and clinical status) and unknown endoscopic remission status, in whom fecal calprotectin, serum CRP, and EHI was measured	Fecal calprotectin <50 μg/g, <150 μg/g, or <250 μg/g Normal CRP (<5 mg/L) EHI <20	Fecal calprotectin >50 μg/g, >150 μg/g, or >250 μg/g Elevated CRP (>5 mg/L) EHI >20	Beneficial: For detection of endoscopic inflammation, TP rate TN rate Harms: FN rate (false reassurance that inflammation has resolved, leading to increased risk of flares due to undertreatment) FP rate (excess endoscopic procedures to rule out inflammation)
Patients with symptomatically active CD					
Question 3	In patients with symptomatically active CD, is an evaluation strategy that combines biomarkers and symptoms superior to symptom-based evaluation for making treatment adjustments?	Patients with symptomatically active CD	Biomarker- and symptom-based evaluation	Symptom-based evaluation	Beneficial: For detection of endoscopic inflammation, TP rate TN rate Harms: FN rate (failure to recognize flare leading to undertreatment/mistreatment, and patient morbidity) FP rate (overdiagnosis, leading to unnecessary treatment adjustment and risk of treatment-related complications)

Table 1. Continued

Variable	Focused question	Patients	Intervention (threshold?)	Comparator	Outcome
Question 4[b]	In patients with symptomatically active CD, at what fecal calprotectin, serum CRP, and EHI cutoffs can we accurately diagnose active inflammation, obviating routine endoscopic assessment?	Patients with established CD with moderate to severe symptoms of CD, or mild symptoms, in whom fecal calprotectin, serum CRP or EHI was measured	Fecal calprotectin >50 μg/g, >150 μg/g, or >250 μg/g Elevated CRP (>5 mg/L) EHI >50	Fecal calprotectin <50 μg/g, <150 μg/g, or <250 μg/g Normal CRP (<5 mg/L) EHI <50	Beneficial: For detection of endoscopic inflammation, TP rate TN rate Harms: FN rate (failure to recognize flare leading to undertreatment/mistreatment, and patient morbidity) FP rate (overdiagnosis, leading to unnecessary treatment adjustment and risk of treatment-related complications)
Patients with CD in surgically induced remission					
Question 5[c]	In patients with CD in surgically induced remission, at what fecal calprotectin, serum CRP, and EHI cutoffs can we accurately rule out postoperative endoscopic recurrence, obviating routine endoscopic assessment?	Asymptomatic patients with CD after surgically induced remission within the past 12 mo, who are at low risk of endoscopic recurrence and on postoperative pharmacologic prophylaxis (10% risk), low risk of endoscopic recurrence and not on postoperative prophylaxis, or at high risk of postoperative recurrence and on postoperative prophylaxis (30% risk), or at high risk of endoscopic recurrence and not on postoperative prophylaxis, who underwent biomarker monitoring 6–12 mo after surgery	Fecal calprotectin <50 μg/g, <150 μg/g, or <250 μg/g Normal CRP (<5 mg/L) EHI <20	Fecal calprotectin >50 μg/g, >150 μg/g, or >250 μg/g Elevated CRP (>5 mg/L) EHI >20	Beneficial: For detection of endoscopic recurrence, TP rate TN rate Harms: FN rate (failure to recognize endoscopic recurrence leading to undertreatment/mistreatment, and patient morbidity) FP rate (overdiagnosis, leading to unnecessary treatment adjustment and risk of treatment-related complications)
Biomarker- vs endoscopy-based monitoring strategy in patients with CD					
Question 9	In patients with established CD, is interval biomarker-based monitoring strategy superior to interval endoscopy-based monitoring strategy to improve long-term outcomes?	Patients with CD in symptomatic remission	Interval biomarker-based monitoring	Interval endoscopy-based monitoring	Maintaining clinical remission at 12 mo and beyond

[a]Test performance of EHI cutoffs for patients with CD in symptomatic remission was addressed in Question 6.
[b]Test performance of EHI cutoffs for patients with symptomatically active CD was addressed in Question 7.
[c]Test performance of EHI cutoffs for patients with CD in surgically induced remission was addressed in Question 8.

undergone endoscopic assessment of disease activity. Patients in whom endoscopic activity status was unknown (ie, assessment of endoscopic activity was more than 3 years ago), we estimated a higher prevalence of endoscopically active inflammation in asymptomatic patients of 45% based on data from clinical trials and observational cohorts. In contrast, in patients with recent confirmation of endoscopic remission without subsequent change in clinical status and therapy were estimated to have a probability of active endoscopic inflammation of 20%, given low likelihood of endoscopic progression in the absence of any change in clinical status and therapy.

- Intermediate pretest probability of having endoscopically active inflammation included patients with mild symptoms of CD (PRO2 8–13, with 3–5 loose or watery stools per day and mild AP, or PRO3 score 13–21). The estimated prevalence of mild to severe endoscopic inflammation in these patients was approximately 65%.
- High pretest probability of having endoscopically active inflammation. These include patients with moderate to severe symptoms of active CD (PRO2 score $>$ 13, with more than 5 loose or watery stools per day and/or moderate to severe AP, or PRO3 score $>$ 21). The estimated prevalence of mild to severe endoscopic inflammation in these patients was approximately 80%.

For detection of postoperative endoscopic recurrence, we examined the performance of each biomarker in detecting significant endoscopic recurrence, defined as Rutgeerts endoscopic score $\geq$i2.[25,26] The panel assumed that initial assessment of endoscopic activity in asymptomatic patients with CD after surgically induced remission would be 6–12 months after the resection. We used 3 illustrative clinical scenarios to determine the likelihood of endoscopic recurrence at this assessment based on individual patient factors influencing risk of postoperative recurrence at time of surgery and use of postoperative pharmacologic prophylaxis. The risk factors include early age at CD diagnosis, smoking, long-segment disease, prior bowel resection, and penetrating disease behavior. Risk of postoperative recurrence is further modified by use of postoperative pharmacologic prophylaxis. Typically, use of postoperative prophylaxis with immunosuppressive therapies lowers the risk of postoperative recurrence by 50%–70%.

- Low risk of postoperative endoscopic recurrence comprised patients without any risk factors associated with a greater likelihood of postoperative recurrence who were on postoperative prophylactic therapy. The estimated likelihood of endoscopic recurrence in this population was approximately 10%.
- Intermediate risk of postoperative endoscopic recurrence comprised patients who had 1 or more risk factors for postoperative recurrence but were on postoperative prophylactic therapy associated with reducing risk of recurrence. The estimated likelihood of endoscopic recurrence in this population was approximately 30%.
- High risk of postoperative endoscopic recurrence comprised patients who had 1 or more risk factors associated with high likelihood of recurrence who were not on postoperative prophylactic therapy. The estimated likelihood of endoscopic recurrence in this population was approximately 60%.

Consequences of Diagnostic Test Results on Patient-Important Outcomes

The panel considered downstream consequences in important patient outcomes corresponding to each possible outcome of a diagnostic test (TP, FP, TN, and FN) (Table 2). Health care providers should be aware of test performance at an individual patient level in each of these scenarios and balance the FN and FP rates with the downstream change in treatment plan that would result from each of these scenarios.

Certainty of the Evidence

We rated the certainty of evidence using the GRADE approach for diagnostic tests and strategies.[16] In this approach, all evidence from RCTs (comparing different diagnostic tests or cutoffs of the same test) and observational diagnostic accuracy studies start at high quality, but can be rated down for any of the following factors:

- risk of bias in included studies (inferred based on QUADAS-2 instrument)[27];
- indirectness (deemed present if there were important differences between the populations studied and those for whom the recommendation is intended); in this updated GRADE approach for diagnostic accuracy studies, TP, FP, TN, and FN derived from sensitivity and specificity were not considered surrogate outcomes;
- inconsistency (deemed present if there were considerable differences between studies in the accuracy estimates that were not explained, or if cutoffs for biomarkers corresponding to endoscopic activity were not prespecified but were obtained primarily post hoc, corresponding to area under the receiver operating characteristic curve);
- imprecision (deemed present if there were wide CIs for TP, FN, TN, and FP rates); and
- publication bias, if strongly suspected.

Evidence profiles were developed for each intervention using the GRADEpro Guideline Development Tool (https://gradepro.org).

Translating Evidence to Recommendations

The guideline panel and evidence synthesis panel met face to face on May 8, 2023 to discuss the evidence and formulate the guideline recommendations. Based on the Evidence-to-Decision framework, the panel considered the certainty of evidence; balance of benefit and harms; patient values and preferences; and, when applicable, feasibility; acceptability; equity; and resource use. For all recommendations, the panel reached consensus. The certainty of evidence and the strength of recommendations are provided for each clinical question. As per GRADE methodology, recommendations are labeled as “strong” or “conditional.” The phrase “we recommend” indicates strong

Table 2. Consequences of Diagnostic Test Results on Patient-Important Outcomes

Test result	Consequence
TPs	Patients correctly diagnosed as having endoscopically active disease would be eligible to undergo treatment adjustment, which may improve symptoms and decrease risk of disease-related complications and morbidity, without being subject to risk, invasiveness, and cost of endoscopic assessment.
FPs	Patients incorrectly diagnosed as having endoscopically active disease, when actually they are in endoscopic remission or have only mild endoscopic activity may undergo unnecessary testing (endoscopy) and/or treatment adjustment, and have avoidable anxiety, potential testing- or treatment-related complications, and increased resource utilization.
TNs	Patients correctly diagnosed as being in endoscopic remission would be reassured and obviate the need for invasive testing with endoscopy, although they may need to undergo serial assessment of biomarker at periodic intervals.
FNs	Patients incorrectly diagnosed as being in endoscopic remission when actually they have moderate to severe endoscopic activity would be falsely reassured, may have avoidable anxiety about unexplained symptoms, and may not receive appropriate treatment adjustment, potentially leading to increased disease related complications, morbidity, and mortality.

recommendations and "we suggest" indicates conditional recommendations. Supplementary Table 2 provides the suggested interpretation of strong and conditional recommendations for patients, clinicians and healthcare policy makers.

Review Process

This guideline was submitted for public comment and peer review. All comments were reviewed and addressed by the full panel. Final recommendations were approved by the AGA Governing Board.

Discussion of Recommendations

All of the recommendations are summarized in Table 3 and are discussed below. Key implementation considerations when contemplating using biomarkers in CD are discussed below and in Table 4.

Key Considerations for Implementing These Recommendations in Clinical Practice

The recommendations in the guideline provide a framework for use of serum or fecal biomarkers in the management of patients with CD to inform treatment. It is important to recognize the limitations of the available data, as well as incorporate both patient and provider thresholds for FP and FN when confronted with a specific clinical setting to implement each guideline recommendation.

1. Considerations of test performance and specificity of biomarkers: Neither serum CRP nor fecal calprotectin are specific for CD activity.
   - CRP may be elevated in systemic inflammatory processes and does not always represent luminal CD activity. Fecal calprotectin is more specific for gut inflammation, but may be elevated in the setting of concomitant gastrointestinal infections. In patients with CD who present with elevated biomarkers and disease-related symptoms, stool testing for *Clostridioides difficile* and other enteric pathogens is important to help rule out other sources of gastrointestinal infections.
2. Role of endoscopic evaluation for other indications: In certain situations, endoscopic assessment may be required for reasons other than assessment of disease activity. Thus, a decision for replacement of endoscopic evaluation by biomarker measurement should consider other information provided by endoscopy.
   - Endoscopic evaluation is warranted for determining the extent and severity of inflammation, dysplasia detection and surveillance, evaluation and endoscopic treatment of stricturing disease, and ruling out cytomegalovirus colitis; biomarkers are not helpful in these situations.
3. Association between treatment target and biomarker performance: The panel debated comparing biomarker performance with any endoscopically active CD (SES-CD $\geq$3) or moderate to severe endoscopic activity (SES-CD $\geq$6). We eventually elected to benchmark biomarker performance against SES-CD $\geq$3, given that this was the threshold used to report biomarker performance in most studies and there was a paucity of data evaluating biomarker performance for detecting moderate to severe endoscopic inflammation (SES-CD $\geq$6). In addition, current CD treatment guidelines recommend a target of endoscopic remission, defined as an SES-CD score $<$3.[20–22] It is likely that biomarker performance would be superior against a higher threshold of SES-CD $>$6. In contrast, the accuracy of biomarkers to detect complete endoscopic healing, defined as an SES-CD of 0, may be less optimal. Yzet et al[28] compared the outcomes of patients with CD who achieved a Crohn's Disease Endoscopic Index of Severity (CDEIS) score of 0 with those of patients with a CDEIS score of 1–4. On longitudinal follow-up, patients with a CDEIS score of 0 had lower rates of treatment failure (25%) compared with those with a score of 1–3 (48%; $P = .047$). None of the patients with a CDEIS score of 0 underwent surgical resection, compared with 11% of patients with a CDEIS score of 1–4. Furthermore, transmural healing, based on cross-sectional imaging or intestinal ultrasound, may also

Table 3. Executive Summary of Recommendations

Recommendation
Patients with CD in symptomatic remission
Recommendation 1: In patients with CD in symptomatic remission, the AGA suggests a monitoring strategy that combines biomarkers and symptoms, rather than relying on symptoms alone. (*Conditional recommendation, low certainty of evidence*) Comment: Patients who place a higher value on avoiding the burden of biomarker testing, over a potentially higher risk of flare and disease progression caused by missing subclinical inflammation, may reasonably choose interval symptom-based monitoring. Implementation considerations: • Interval biomarker monitoring may be performed every 6–12 mo in patients in symptomatic remission. • Biomarker-based monitoring may be particularly useful in patients w biomarkers have historically correlated with endoscopic disease activity.
Recommendation 2: In patients with CD in symptomatic remission with recent confirmation of endoscopic remission (without any change in clinical status, on stable therapy), the AGA suggests using fecal calprotectin <150 μg/g and/or CRP <5 mg/L to rule out active inflammation, and avoid routine endoscopic assessment of disease activity. (*Conditional recommendation, low to moderate certainty of evidence*)
Recommendation 3: In patients with CD in symptomatic remission without recent confirmation of endoscopic remission, the AGA suggests endoscopic evaluation to rule out active inflammation, rather than relying solely on fecal calprotectin or CRP. (*Conditional recommendation, low to moderate certainty of evidence*) Implementation considerations: • The panel considered recent confirmation of endoscopic or radiologic remission to ideally have been within 3 y. • Radiologic assessment of disease activity may be a reasonable alternative to endoscopic assessment for patients with predominantly small bowel involvement.
Recommendation 4: In patients with CD in symptomatic remission, with elevated biomarkers of inflammation (fecal calprotectin >150 μg/g, CRP >5 mg/L), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, low certainty of evidence*) Implementation considerations: • In patients with CD in sustained symptomatic remission but elevated biomarkers, repeat measurement of biomarkers (in 3–6 mo) may be a reasonable alternative to endoscopic (or radiologic) assessment, especially if the latter has been performed recently. • Lack of normalization of biomarkers (or persistently elevated biomarkers) in patients whose symptoms recently resolved after initial treatment of symptomatically active CD, likely suggests active inflammation, and may warrant treatment adjustment, without need for endoscopic (or radiologic) evaluation.
Patients with symptomatically active CD
Recommendation 5: In patients with symptomatically active CD, the AGA suggests a biomarker-based assessment and treatment adjustment strategy, rather than relying on symptoms alone. (*Conditional recommendation, moderate certainty of evidence*) Comment: Patients who place a higher value on avoiding the burden of biomarker testing, over a potentially higher risk of over- or undertreatment if relying only on symptoms, may consider choosing interval symptom-based treatment adjustment when being treated for active symptoms. Implementation considerations: • Interval biomarker assessment and treatment adjustment may be performed every 2–4 mo in patients being treated for active symptoms. • After resolution of symptoms (and normalization of biomarkers), endoscopic (and/or radiologic) evaluation should be performed to rule out active inflammation, typically 6–12 mo after treatment initiation or adjustment. The patient may then transition to guidance for patients in symptomatic remission.
Recommendation 6: In patients with CD with mild symptoms and elevated biomarkers of inflammation (fecal calprotectin >150 μg/g, CRP >5 mg/L), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, very low certainty of evidence*) Implementation consideration: • Lack of normalization of biomarkers (or persistently elevated biomarkers) in patients whose symptoms partially improve after initial treatment of active CD, likely suggests active inflammation, and may warrant treatment adjustment, without need for endoscopic (or radiologic) evaluation.
Recommendation 7: In patients with CD with mild symptoms and normal biomarkers of inflammation (fecal calprotectin <150 μg/g, CRP <5 mg/L), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, very low certainty of evidence*)

Table 3. Continued

Recommendation
Recommendation 8. In patients with CD with moderate to severe symptoms, the AGA suggests using fecal calprotectin >150 μg/g or CRP >5 mg/L, to rule in active inflammation and inform treatment adjustment and avoid routine endoscopic assessment of disease activity. (*Conditional recommendation, low to moderate certainty of evidence*)
Recommendation 9: In patients with CD with moderate to severe symptoms with normal biomarkers of inflammation (fecal calprotectin <150 μg/g, CRP <5 mg/L), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, low certainty of evidence*)
Patients with CD in surgically induced remission
Recommendation 10: In asymptomatic patients with CD after surgically induced remission within the past 12 mo, who are at low risk of postoperative recurrence or who have 1 or more risk factors for recurrence but are on postoperative pharmacologic prophylaxis, the AGA suggests using fecal calprotectin <50 μg/g, to avoid routine endoscopic assessment of disease activity. (*Conditional recommendation, moderate certainty of evidence*)
Comment: Patients, particularly those with multiple prior surgeries, and/or with failure of multiple advanced therapies before surgery, who value more accurate assessment of endoscopic recurrence over the inconvenience and costs of colonoscopy, may reasonably choose endoscopic assessment of disease activity within 12 mo after surgery.
Recommendation 11: In asymptomatic patients with CD after surgically induced remission within the past 12 mo, who are at high baseline risk of recurrence and are not receiving postoperative pharmacologic prophylaxis, the AGA suggests endoscopic evaluation, rather than relying solely on biomarkers, for assessing endoscopic recurrence. (*Conditional recommendation, low to moderate certainty of evidence*)
Implementation considerations: • Risk stratification schemes to classify patients' risk of endoscopic recurrence after surgically induced remission are not well-defined. Risks factors typically associated with low risk of recurrence 6–12 mo after surgically induced remission include older age at surgery (older than 50 y), nonsmoking, long-standing disease (more than 10 y), and first surgery for a short segment of fibrostenotic disease (<10–20 cm). Risk factors typically associated with high risk of recurrence 6–12 mo after surgically induced remission include 2 or more prior surgeries, penetrating or perianal disease, smoking, young age at surgery, with long segment of small bowel resection; however, these risk factors may not be additive. • In patients at low baseline risk of recurrence, who are also receiving postoperative pharmacologic prophylaxis, fecal calprotectin <150 μg/g may also rule out endoscopic recurrence. • Normal CRP in patients with asymptomatic CD in surgically induced remission is not able to rule out endoscopic recurrence accurately. • There are limited data on ongoing biomarker monitoring alone in patients with CD in surgically induced remission. Colonoscopic evaluation may be warranted beyond 12 mo after surgery in patients when biomarker-based monitoring is being pursued.
EHI (Monitr) in patients with CD
Recommendation 12: In patients with CD, the AGA suggests neither in favor of nor against the use of EHI (Monitr) for monitoring inflammation and treatment decisions. (*No recommendation, knowledge gap*)
Biomarker- vs endoscopy-based monitoring strategy in patients with CD
Recommendation 13: In patients with CD, the AGA makes no recommendation in favor of, or against, a biomarker-based monitoring strategy over an endoscopy-based monitoring strategy to improve long-term outcomes. (*No recommendation, knowledge gap*)

have prognostic significance in patients with CD. Diagnostic performance of a combination of symptoms and biomarkers to detect these more rigorous end points was not assessed in this guideline, but is likely to have inferior performance, given differences in pretest probability. In this guideline, we focused on the accuracy of biomarkers to detect active endoscopic inflammation (mild to severe activity), defined as the absence of endoscopic remission to be consistent with clinical practice and RCT end points.

- Test performance of all biomarkers in this guideline reflects their ability to rule out active endoscopic inflammation (SES-CD >3). Biomarkers may be suboptimal for detecting more rigorous treatment targets, such as SES-CD of 0, transmural healing, or histologic remission. However, biomarker performance may be better than reported in the guidelines for detecting moderate to severe endoscopic activity (SES-CD >6).

4. Influence of disease location on performance of fecal biomarkers and correlation with symptoms: Elevation of fecal calprotectin may be influenced by the extent and location of inflamed surface. The panel identified limited data comparing the performance of fecal calprotectin by disease location. A systematic review by Simon et al[29] identified 16 eligible studies that examined the sensitivity and specificity of fecal calprotectin by disease location. The sensitivity of fecal calprotectin for small bowel inflammation ranged from 43% to 100%, and that for large bowel disease ranged from 67% to 100%. Three studies noted fecal calprotectin correlated with endoscopic severity in the large bowel alone, and 2 other

Table 4. Key Considerations When Using Biomarkers for Monitoring in Crohn's Disease

Key considerations
1. Considerations of test performance and specificity of biomarkers: CRP may be elevated in systemic inflammatory processes and does not always represent luminal CD activity. Fecal calprotectin is more specific for gut inflammation, but may be elevated in the setting of concomitant gastrointestinal infections. In patients with CD who present with elevated biomarkers and disease-related symptoms, stool testing for *Clostridioides difficile* and other enteric pathogens is important to help rule out other sources of gastrointestinal infections.
2. Role of endoscopic evaluation for other indications: Endoscopic evaluation is warranted for determining the extent and severity of inflammation, dysplasia detection, and surveillance, evaluation and endoscopic treatment of stricturing disease, and ruling out cytomegalovirus colitis; biomarkers are not helpful in these situations.
3. Association between treatment target and biomarker performance: Test performance of all biomarkers in this guideline reflects their ability to rule out active endoscopic inflammation (SES-CD >3). Biomarkers may be suboptimal for detecting more rigorous treatment targets, such as SES-CD of 0, transmural healing, or histologic remission. However, biomarker performance may be better than reported in the guidelines to detect moderate to severe endoscopic activity (SES-CD >6).
4. Influence of disease location on performance of fecal biomarkers and correlation with symptoms: Gastrointestinal symptoms may correlate less accurately with endoscopic activity in patients with small bowel CD or those with prior intestinal resection compared with patients with predominant or extensive colonic involvement. Fecal biomarkers may be modestly less accurate in detecting endoscopic inflammation in small bowel CD or upper gastrointestinal disease than patients with predominant or extensive colonic involvement. In order to interpret results of fecal biomarkers in patients with predominantly small bowel involvement, close anchoring of symptoms and biomarkers with endoscopic findings (ie, measuring biomarker and endoscopic activity simultaneously) in patients with active disease, and in remission, is preferred.
5. Interpreting biomarker performance for low-risk vs high-risk treatment adjustments: Application of all biomarkers in clinical practice should be guided by downstream implications, including risk of consequent treatment decisions (low-risk treatment adjustment vs high-risk treatment adjustment). Test performance thresholds (acceptable FP and FN rates) may vary for patient–provider teams, depending on what treatment adjustment is being considered.
6. Inter- and intra-assay test variability: Fecal calprotectin assays may not be interchangeable and the same assay should be used for a given patient to compare results over time. Because there can be substantial within-stool and within-day variation of fecal calprotectin measurements from a single patient, confidence in any single measurement may be limited. Hence, if there is uncertainty of results (such as borderline or unexpected results), repeat fecal calprotectin testing or endoscopic evaluation for confirmation may be required.
7. Inter-individual heterogeneity in biomarkers responsiveness: There are inter-individual differences in biomarker elevation in patients with intestinal inflammation, and in a subset of patients, biomarkers may correlate poorly with endoscopic activity. The overall performance and confidence in the use of biomarkers for treatment decisions in a particular patient may be higher when these biomarkers have been longitudinally observed to correlate with the patient's endoscopic disease activity (both during active disease and remission).

studies demonstrated similar correlation in both small bowel and large bowel disease locations. From 11 studies that compared performance across disease locations, 4 studies demonstrated superior performance in large bowel CD and 7 other studies found no difference in performance between small bowel and large bowel locations. Similarly, studies have suggested that the correlation between symptoms and endoscopic activity may be influenced by disease location and prior bowel surgery, with a stronger correlation between symptoms and endoscopic activity observed in patients with CD with colon-dominant disease vs patients with predominantly small bowel involvement, and in patients without prior intestinal resection.[10–12] We were unable to critically analyze the diagnostic performance of fecal calprotectin in various clinical scenarios by disease location and, hence, opted to report diagnostic performance for small bowel and colonic CD together. The guideline panel posited that, beyond disease location, the extent and severity of involved segments may have a considerable impact on test performance, independent of disease location, similar to observations in patients with UC, when the performance of fecal calprotectin may be inferior in patients with limited proctitis vs extensive colitis. Unfortunately, studies did not report the performance of fecal calprotectin by disease extent, separate from disease location.

- Gastrointestinal symptoms may correlate less accurately with endoscopic activity in patients with small bowel CD or those with prior intestinal resection compared with patients with predominant or extensive colonic involvement.
- Fecal biomarkers may be modestly less accurate in detecting endoscopic inflammation in small bowel CD or upper gastrointestinal disease than patients with predominant or extensive colonic involvement. In order to interpret results of fecal biomarkers in patients with predominantly small bowel involvement, close anchoring of symptoms and biomarkers with endoscopic findings (ie, measuring biomarker and endoscopic activity simultaneously) in patients with active disease, and in remission, is preferred.

5. Interpreting biomarker performance for low-risk vs high-risk treatment adjustments: The acceptable threshold for

performance of biomarkers may differ based on the absolute and/or perceived cost and risk of the proposed interventions in response to biomarker thresholds. For example, in patients with CD with symptoms, a higher rate of FP (ie, patients incorrectly labeled as having active endoscopic inflammation based on biomarkers, when their disease is actually in remission) may be acceptable for lower risk treatment adjustments, such as a brief course of steroids in individuals at low risk for adverse effects. However, it is reasonable to accept lower FP rates for interventions that may be associated with significant cost (dose escalation of biologic therapy) or risk (change in therapy).

- Application of all biomarkers in clinical practice should be guided by downstream implications, including risk of consequent treatment decisions (low-risk treatment adjustment vs high-risk treatment adjustment). Test performance thresholds (acceptable FP and FN rates) may vary for patient–provider teams, depending on what treatment adjustment is being considered.

6. Inter- and intra-assay test variability: Biomarker levels may vary between laboratories. Thus, use of the same assay type and laboratory are preferred for accurate comparison of biomarker trajectory.

   - Fecal calprotectin assays may not be interchangeable and the same assay should be used for a given patient to compare results over time. Because there can be substantial within-stool and within-day variation of fecal calprotectin measurements from a single patient, confidence in any single measurement may be limited. Hence, if there is uncertainty of results (such as borderline or unexpected results), repeat fecal calprotectin testing or endoscopic evaluation for confirmation may be required.

7. Inter-individual heterogeneity in biomarker responsiveness: Biomarkers including CRP and fecal calprotectin demonstrate heterogeneity between individuals. Up to one-fifth of patients may not demonstrate an elevation in these biomarkers in the setting of endoscopically active disease. Consequently, it is important to anchor the performance of a biomarker against endoscopic assessment for a given patient both in active disease and in remission. Biomarker accuracy is likely superior and of greater clinical value in a patient where the biomarker was shown to be elevated in the setting of endoscopically active disease and normalizes with resolution of inflammation. In patients where this correlation has not been demonstrated before, interpretation of biomarker result may merit more caution.

   - There are inter-individual differences in biomarker elevation in patients with intestinal inflammation, and in a subset of patients, biomarkers may correlate poorly with endoscopic activity. The overall performance and confidence in the use of biomarkers for treatment decisions in a particular patient may be higher when these biomarkers have been longitudinally observed to correlate with the patient's endoscopic disease activity (both during active disease and remission).

Guideline Recommendations

Patients With Crohn's Disease in Symptomatic Remission

GUIDELINES

Question 1: In patients with CD in symptomatic remission, is interval biomarker-based monitoring superior to symptom-based monitoring to improve long-term outcomes?

Recommendation 1: In patients with CD in symptomatic remission, the AGA suggests a monitoring strategy that combines biomarkers and symptoms, rather than relying on symptoms alone. (*Conditional recommendation, low certainty in evidence*)

Comment: Patients who place a higher value on avoiding the burden of biomarker testing, over a potentially higher risk of flare and disease progression caused by missing subclinical inflammation, may reasonably choose interval symptom-based monitoring.

Implementation considerations:

- Interval biomarker monitoring may be performed every 6–12 months in patients in symptomatic remission.
- Biomarker-based monitoring may be particularly useful in patients where biomarkers have historically correlated with endoscopic disease activity.

Summary of the Evidence

The panel compared a biomarker-based monitoring strategy with routine and systematic checking of biomarkers against monitoring of symptoms alone to guide treatment changes in patients with established CD in symptomatic remission. Supplementary Figure 3 lays out the schematic for the proposed comparison. We did not identify any RCTs that directly compared the 2 strategies and could inform our recommendations. Although the CALM trial compared a symptom-based treatment adjustment strategy with biomarker-based treatment adjustment, all participants had active disease at study entry.[14]

Similarly, the STARDUST (Study of Treat to Target Versus Routine Care Maintenance Strategies in Crohn's Disease Patients Treated With Ustekinumab) trial compared symptom-based vs symptom- and biomarker-based treatment escalation with early endoscopic assessment to guide increases in dose of ustekinumab, but all patients were symptomatically active at study entry.[30] Thus, both studies did not directly inform this study question. We subsequently examined cohort studies in patients with CD in symptomatic remission, comparing rates of disease relapse over long-term follow-up between those with elevated and normal biomarkers. A meaningfully higher risk of relapse in those with elevated biomarkers in symptomatic remission would support a biomarker-based monitoring strategy in CD. We identified 12 cohort studies comprising 982 patients in symptomatic remission with unknown endoscopic

activity at enrollment. All of these studies examined fecal calprotectin as the biomarker. One-third of patients (38%) had elevated fecal calprotectin, defined variably as >200–300 μg/g (Supplementary Figure 4). At median follow-up of 1 year, patients with elevated fecal calprotectin were 4.8 times more likely to have disease relapse compared with patients with normal fecal calprotectin (95% CI, 2.81–8.17), with a high degree of heterogeneity (I^2 = 82%). With an observed median annual risk of relapse of 11% in patients with CD in symptomatic remission and normal fecal calprotectin, the risk of relapse over 12 months in those with elevated biomarkers and symptomatic remission was 52.7% (95% CI, 30.9%–89.9%) (Table 5).

Benefits and Harms (Downsides)

Symptom-based monitoring strategy. The benefit of a symptom-based monitoring strategy is that it relies on symptoms usually assessed as part of routine ongoing clinical care. However, the potential harms of this strategy would be a potentially higher risk of relapse due to potentially missing ongoing endoscopically active and clinical meaningful inflammation in asymptomatic individuals.

Biomarker-based monitoring strategy. The benefit of a biomarker-based strategy may be a more accurate prognostication of disease outcomes over 1 year by identifying individuals who may have ongoing endoscopically active disease, despite the absence of symptoms, potentially allowing for early treatment adjustment before symptomatic relapse. The potential harms of a biomarker-based strategy are the cost and inconvenience, particularly for stool-based biomarkers. Elevated biomarkers in asymptomatic individuals may also lead to anxiety and increased costs due to need for downstream testing to determine FP rates (see Question 2, Recommendation 4).

Certainty of Evidence

When examining cohort studies comparing long-term outcomes in patients with CD in symptomatic remission with elevated vs normal biomarkers, there was low confidence in effect estimates supporting the use of a biomarker-based monitoring strategy over a symptom-based monitoring strategy. Evidence was rated down for risk of bias in included studies and inconsistency in effect estimates with variability in cutoffs of fecal calprotectin. There were limited data on the prognostic value of other biomarkers, such as serum CRP, in patients with asymptomatic CD.

Rationale

Using the GRADE Evidence-to-Decision framework, incorporating the potential benefits and downsides of both strategies, the guideline panel conditionally recommended a biomarker-based monitoring strategy over symptom-based monitoring alone. Some patients who prefer to avoid the burden of biomarker-based monitoring in terms of cost and inconvenience may reasonably decide to adopt a symptom-based monitoring strategy alone. The panel determined that an interval of 6–12 months for monitoring biomarkers would be reasonable to reflect routine clinic follow-ups for most patients with CD. As biomarkers may not perform equally well in all patients, the recommendation for biomarker-based monitoring is best suited to those who have previously demonstrated a good correlation between their endoscopic inflammation and biomarker elevation. The panel could not identify literature supporting the efficacy of downstream treatment adjustments in response to biomarker elevation alone in asymptomatic individuals, particularly those in sustained remission. The panel acknowledged that it may be reasonable, in a subset of patients, to follow-up an elevated biomarker measure with serial monitoring rather than treatment escalation or immediate endoscopic assessment. The literature was inadequate to examine the relative prognostic value of different cutoffs for elevated biomarkers, and the panel acknowledged that the prognostic value of a markedly elevated biomarker may differ from mild elevation.

Question 2: In patients with CD in symptomatic remission, at what fecal calprotectin, serum C-reactive protein, and Endoscopic Healing Index cutoff can we accurately rule out active inflammation, obviating routine endoscopic assessment?

Recommendation 2: In patients with CD in symptomatic remission with recent confirmation of endoscopic remission (without any change in clinical status, on stable therapy), the AGA suggests using fecal calprotectin <150 μg/g and/or CRP <5 mg/L (or below cutoff for normal range for the laboratory) to rule out active inflammation, and avoid routine endoscopic assessment of disease activity. (*Conditional recommendation, low to moderate certainty in evidence*)

Recommendation 3: In patients with CD in symptomatic remission without recent confirmation of endoscopic remission, the AGA suggests endoscopic evaluation to rule out active inflammation, rather than relying solely on fecal calprotectin or CRP. (*Conditional recommendation, low to moderate certainty in evidence*)

Implementation considerations:

- The panel considered recent confirmation of endoscopic or radiologic remission to ideally have been within 3 years.
- Radiologic assessment of disease activity may be a reasonable alternative to endoscopic assessment for patients with predominantly small bowel involvement.

Recommendation 4: In patients with CD in symptomatic remission, with elevated biomarkers of inflammation (fecal calprotectin >150 μg/g, CRP >5 mg/L), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, low certainty in evidence*)

Implementation considerations:

- In patients with CD in sustained symptomatic remission but elevated biomarkers, repeat measurement of

Table 5. Grading of Recommendations Assessment, Development and Evaluation Evidence Profile for PICO (Patients, Intervention, Comparator, and Outcome) Question 1, Comparing Outcomes With Interval Biomarker-Based Monitoring vs Symptom-Based Monitoring to Improve Long-Term Outcomes in Patients With Crohn's Disease in Symptomatic Remission[a]

Outcome No. of participants (studies)	Relative effect (95% CI)	Anticipated absolute effects (95% CI)			Quality
		Normal fecal calprotectin	Elevated fecal calprotectin	Difference	
Risk of relapse at 12 mo 982 (12 cohort)	RR, 4.79 (2.81–8.17)	Pooled relapse rate at 1 y, %			⨁⨁◯◯ LOW[b,c]
		11	52.7 (30.9–89.9)	41.7 more (19.9 more to 78.9 more)	

RR, relative risk.
[a]Patient or population: patients with CD in symptomatic remission; setting: cohort; exposure: elevated fecal calprotectin (>200–300 μg/g); comparison: normal fecal calprotectin.
[b]Risk of bias based on Quality in Prognosis Studies tool.
[c]Inconsistency in effect estimates; variability in fecal calprotectin cutoffs.

biomarkers (in 3–6 months) may be a reasonable alternative to endoscopic (or radiologic) assessment, especially if the latter has been performed recently.

- Lack of normalization of biomarkers (or persistently elevated biomarkers) in patients whose symptoms recently resolved after initial treatment of symptomatically active CD, likely suggests active inflammation, and may warrant treatment adjustment, without need for endoscopic (or radiologic) evaluation.

Summary of the Evidence

Diagnostic performance of fecal calprotectin. The evidence synthesis team examined the following 3 diagnostic cutoffs for fecal calprotectin: 50 ± 50 μg/g, 150 ± 50 μg/g, and 250 ± 50 μg/g through a systematic review and meta-analysis of published studies (Table 6). To minimize bias due to selective reporting of optimized cutoffs (as is common in diagnostic accuracy studies), we included only studies that reported diagnostic accuracy of preselected fecal calprotectin cutoffs or reported the performance across 2 or more predetermined cutoffs. The gold standard for comparison was either endoscopically active CD (SES-CD score ≥3) or endoscopic remission (SES-CD <3). The sensitivity and specificity of fecal calprotectin cutoff of 50 ± 50 μg/g were 88% (95% CI, 79%–94%) and 67% (95% CI, 51%–80%), respectively, based on 16 cohorts. The corresponding sensitivity and specificity of 150 ± 50 μg/g cutoff (11 cohorts) were 81% (95% CI, 74%–87%) and 72% (95% CI, 61%–81%), respectively, and of 250 ± 50 μg/g cutoff (14 cohorts) were 76% (95% CI, 70%–82%) and 74% (95% CI, 67%–80%), respectively (Supplementary Figure 5).

Test performance in patients with Crohn's disease with known recent endoscopic remission within preceding 3 years (with 20% prevalence of active inflammation in asymptomatic patients). We estimated that, among patients with CD in endoscopic remission on stable therapy and without any change in clinical status, <20% will have progression to endoscopic inflammation over 2–3 years. In applying fecal calprotectin performance cutoffs to this scenario, approximately 2.4%, 3.8%, and 4.8% of patients (FN rates) with fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g, respectively, may be misclassified as having endoscopic remission when they actually have endoscopically active inflammation (Table 6). In contrast, elevated fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g in asymptomatic patients had significantly elevated FP rates of 26.4%, 22.4%, and 20.8%, whereby patients may be misclassified as having endoscopically active disease, despite being in endoscopic remission.

Test performance in patients with Crohn's disease with unknown endoscopic remission status (with 45% prevalence of endoscopically active inflammation). In this scenario, approximately 5.4%, 8.5%, and 10.8% of patients (FN rate) with fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g, respectively, may be misclassified as having endoscopic remission when they have endoscopically active disease (Table 6). In contrast, elevated fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g also had high FP rates of 18.1%, 15.4%, and 14.3%, that is, a substantial proportion of patients with endoscopic remission may be incorrectly classified as having endoscopic activity.

Diagnostic performance of serum C-reactive protein. We identified 20 studies reporting on the performance of CRP in this setting; most studies used a cutoff of >5 mg/L. The sensitivity of elevated CRP to detect endoscopically active disease was 67% (95% CI, 54%–77%) with a specificity of 73% (95% CI, 65%–80%) (Supplementary Figure 6).

Known recent endoscopic remission within preceding 3 years (20% prevalence of endoscopic activity). In applying this cutoff (elevated CRP, generally >5 mg/L), approximately 6.6% of patients would have an FN result and be mislabeled as being in endoscopic remission while having endoscopically active disease (Table 7). Elevated CRP (>5 mg/L) in this setting had an FP rate of 21.6%, suggesting that nearly one-fourth of patients may be mislabeled as having endoscopic active disease while being in remission.

Table 6. Grading of Recommendations Assessment, Development and Evaluation Evidence Profile for PICO (Patients, Intervention, Comparator, and Outcome) Question 2, Comparing Cutoffs for Fecal Calprotectin in Patients With Crohn's Disease Without Symptoms With Known Endoscopic Remission or With Unknown Endoscopic Remission Status[a]

	No. of results per 1000 patients tested (95% CI)						
	Known endoscopic remission (prevalence 20%)			Unknown endoscopic remission status (prevalence 45%)			
Test result	fCal <50 μg/g	fCal <150 μg/g	fCal <250 μg/g	fCal <50 μg/g	fCal <150 μg/g	fCal <250 μg/g	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	176 (158–188)	162 (148–174)	152 (140–164)	396 (356–423)	365 (333–392)	342 (315–369)	TPs would be eligible to undergo treatment adjustment, which may decrease disease-related complications and morbidity, without being subject to risks and invasive testing with endoscopy.
FNs (patients incorrectly labeled as being in endoscopic remission or having only mild endoscopic activity, when actually they have moderate to severe endoscopic activity)	24 (12–42)	38 (26–52)	48 (36–60)	54 (27–94)	85 (58–117)	108 (81–135)	FNs would be falsely reassured, and may be at higher risk of disease complications/flare due to undertreatment.
GRADE certainty of evidence	MODERATE[b]	LOW[b,c]	LOW[b,c]	LOW[b,c]	VERY LOW[b,d]	VERY LOW[b,d]	
TNs (patients correctly diagnosed as being in endoscopic remission or having only mild endoscopic activity)	536 (408–640)	576 (488–648)	592 (536–640)	369 (281–440)	396 (336–446)	407 (369–440)	TNs would be reassured and obviate the need for invasive testing with endoscopy, although they may need to undergo serial assessment of biomarker at periodic intervals.

Table 6. Continued

Test result	No. of results per 1000 patients tested (95% CI)						Comments
	Known endoscopic remission (prevalence 20%)			Unknown endoscopic remission status (prevalence 45%)			
	fCal <50 μg/g	fCal <150 μg/g	fCal <250 μg/g	fCal <50 μg/g	fCal <150 μg/g	fCal <250 μg/g	
FPs (patients incorrectly labeled as having moderate to severe endoscopic activity, when actually they are in endoscopic remission or have only mild endoscopic activity)	264 (160–392)	224 (152– 312)	208 (160–264)	181 (110–269)	154 (104–214)	143 (110–181)	FPs may receive unnecessary testing (endoscopy) and/or treatment adjustment, and have avoidable anxiety, potential testing- or treatment-related complications and excessive resource utilization.
GRADE certainty of evidence	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	—

fCal, fecal calprotectin.
[a]Population/setting: Adults with CD in symptomatic remission on stable maintenance therapy, with known endoscopic remission asymptomatic + known endoscopic remission in preceding 2–3 y, without change in clinical status and on stable therapy with observed prevalence of active disease of 20%; with unknown endoscopic remission status (asymptomatic + unknown endoscopic remission status in preceding 2–3 y) with observed prevalence of endoscopically active disease of 45%. Pooled sensitivity/specificity fCal with cutoff <50 μg/g: sensitivity, 88% (95% CI, 79%–94%); specificity, 67% (95% CI, 51%–80%), 16 studies. Pooled sensitivity/specificity fCal with cutoff <150 μg/g: sensitivity, 81% (95% CI, 74%–87%); specificity, 72% (95% CI, 61%–81%), 13 studies. Pooled sensitivity/specificity fCal with cutoff <250 μg/g: sensitivity, 76% (9% CI, 70%–82%); specificity, 74% (95% CI, 67%–80%), 14 studies. Reference test: colonoscopy.
[b]High unexplained heterogeneity, selective inclusion of studies reporting cutoffs.
[c]Serious imprecision because 95% CI crosses maximal tolerable FN threshold of <5%.
[d]Very serious imprecision because point estimate and 95% CI are higher than maximal tolerable FN threshold.
[e]Very serious imprecision because point estimate is higher than maximal tolerable FP threshold.

Table 7. Grading of Recommendations Assessment, Development and Evaluation Evidence Profile for PICO (Patients, Intervention, Comparator, and Outcome) Question 2, Comparing Cutoffs for Serum C-Reactive Protein in Patients With Crohn's Disease Without Symptoms With Known Endoscopic Remission or With Unknown Endoscopic Remission Status[a]

	No. of results per 1000 patients tested (95% CI)		
	Known endoscopic remission (prevalence 20%)	Unknown endoscopic remission status (prevalence 45%)	
Test result	Normal CRP	Normal CRP	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	134 (108–154)	302 (243–347)	TPs would be eligible to undergo treatment adjustment, which may decrease disease-related complications and morbidity, without being subject to risks and invasive testing with endoscopy.
FNs (patients incorrectly labeled as being in endoscopic remission or having only mild endoscopic activity, when actually they have moderate to severe endoscopic activity)	66 (46–92)	148 (103–207)	FNs would be falsely reassured, and may be at higher risk of disease complications/flare due to undertreatment.
GRADE Certainty of evidence	LOW[b,c]	VERY LOW[b,d]	
TNs (patients correctly diagnosed as being in endoscopic remission or having only mild endoscopic activity)	584 (520–640)	402 (358–440)	TNs would be reassured and obviate the need for invasive testing with endoscopy, although they may need to undergo serial assessment of biomarker at periodic intervals.
FPs (patients incorrectly labeled as having moderate to severe endoscopic activity, when actually they are in endoscopic remission or have only mild endoscopic activity)	216 (160–280)	148 (110–192)	FPs may receive unnecessary testing (endoscopy) and/or treatment adjustment, and have avoidable anxiety, potential testing- or treatment-related complications and excessive resource utilization.
GRADE certainty of evidence	VERY LOW[b,e]	VERY LOW[b,e]	

[a]Population/setting: adults with CD in symptomatic remission on stable maintenance therapy, with known endoscopic remission asymptomatic + known endoscopic remission in preceding 2–3 y, without change in clinical status and on stable therapy) with observed prevalence of active disease of 20%; with unknown endoscopic remission status (asymptomatic + unknown endoscopic remission status in preceding 2–3 y) with observed prevalence of endoscopically active disease of 45%. Pooled sensitivity of CRP <5 mg/L, 66.7% (95% CI, 54.4–77.1), 20 studies. Pooled specificity of CRP <5 mg/L, 73.1% (95% CI, 64.7–80.1), 20 studies. Reference test: colonoscopy.
[b]High unexplained heterogeneity, selective inclusion of studies reporting cutoffs.
[c]Serious imprecision because 95% CI crosses maximal tolerable FN threshold of <5%.
[d]Very serious imprecision because point estimate and 95% CI are higher than maximal tolerable FN threshold.
[e]Very serious imprecision because point estimate and 95% CI are higher than maximal tolerable FP threshold.

Unknown endoscopic remission or endoscopic remission confirmation more than 3 years ago (45% prevalence of endoscopically active inflammation). In this scenario, 14.8% of patients would have an FN result and be mislabeled as being in remission while having endoscopically active disease. A similar proportion (14.8%) would have an FP result and be wrongly characterized as having endoscopically active disease while in remission.

Certainty of Evidence

There was no direct evidence comparing how different biomarker cutoffs and accompanying treatment decisions impact downstream patient-important outcomes. However, we did not rate down for indirectness because the presence of endoscopic activity is a close surrogate for unfavorable patient outcomes, and an indication for treatment adjustment.

Fecal calprotectin. There was moderate certainty of evidence supporting the use of fecal calprotectin cutoffs of <50 μg/g (evidence rated down for inconsistency due to selective inclusion of studies reporting specific cutoffs and high heterogeneity for summary sensitivity and specificity), and low certainty of evidence supporting the use of fecal calprotectin cutoffs of <150 μg/g and <250 μg/g (evidence rated down for inconsistency and imprecision because 95% CI exceeded the maximal tolerable FN rate of 5%) to rule out endoscopic inflammation in patients with known endoscopic remission. In patients with unknown endoscopic remission status, the corresponding certainty of evidence is low for calprotectin <50 μg/g and very low for cutoffs of <150 μg/g and <250 μg/g (evidence rated down for heterogeneity and very serious imprecision because both the point estimate and 95% CI are higher than FN threshold of 5%). In patients with CD in symptomatic remission, either with known endoscopic remission or unknown endoscopic remission status, there was very low certainty of evidence supporting the use of any proposed cutoff of elevated fecal calprotectin to rule in endoscopic inflammation, due to unacceptably high rates of FP (very serious imprecision) and inconsistency.

Serum C-reactive protein. There was low certainty of evidence supporting the use of CRP <5 mg/L to rule out endoscopic inflammation in patients with CD in symptomatic remission and with known endoscopic remission. Evidence was rated down for inconsistency due to selective reporting of cutoffs in studies optimized for best performance and high heterogeneity for summary sensitivity and specificity, and for serious imprecision because 95% CI exceeded the maximal tolerable FN rate of 5%. In patients in whom endoscopic remission status was unknown, evidence supporting the use of CRP <5 mg/L to rule out endoscopic inflammation was very low due to unacceptably high rates of FN (very serious imprecision) and inconsistency. In both endoscopic remission scenarios, there was very low certainty of evidence supporting the use of elevated CRP to rule in endoscopic inflammation, due to unacceptably high rates of FP (very serious imprecision) and inconsistency.

Rationale

For the appropriate use of biomarkers in the assessment of patients with CD, patients and health care providers should incorporate both test performance and the downstream consequence of FP and FN rates. The panel acknowledged that there may be instances when patients and providers may be willing to accept higher (>5%) FN rates, depending on the downstream consequences. For ease of implementation in clinical practice and for consistency with UC guidelines,[15] the guideline panel preferred choosing a single fecal calprotectin cutoff (<150 μg/g) that is applicable to multiple scenarios. However, there may be specific clinical situations when a higher or lower cutoff may have an acceptable performance, such as cutoff of fecal calprotectin <250 μg/g in asymptomatic patients with known endoscopic remission. Conceivably, in patients with small bowel–dominant CD, where correlation between symptoms and endoscopic activity is less strong and performance of fecal calprotectin may be modestly lower, lower fecal calprotectin thresholds, such as <150 μg/g or <50 μg/g, may yield lower FN rates.

There were limited data regarding the predictive value of serially measured biomarkers. In asymptomatic patients with elevated biomarkers, the FP rate for elevated fecal calprotectin or CRP was sufficiently high that the panel recommended endoscopic assessment before treatment adjustment to minimize likelihood of overtreatment. However, in this clinical scenario, the panel also recognized that it may be reasonable, in some patients, to consider serial monitoring of biomarkers and determine the trajectory of elevation as a factor in informing downstream actions. There were also insufficient data to inform examination of combinations of biomarkers. For example, does the presence of an elevated fecal calprotectin and an elevated CRP increase the likelihood of endoscopically active disease beyond elevation of either biomarker alone? In general, the sensitivity of fecal calprotectin was greater when compared with CRP. However, CRP may be more readily measured and integrated into routine clinical practice.

Patients With Symptomatically Active Crohn's Disease

Question 3: In patients with symptomatically active CD, is an evaluation strategy that combines biomarkers and symptoms superior to symptom-based evaluation for making treatment adjustments?

Recommendation 5: In patients with symptomatically active CD, the AGA suggests a biomarker-based assessment and treatment adjustment strategy, rather than relying on symptoms alone. (*Conditional recommendation, moderate certainty in evidence*)

Comment: Patients who place a higher value on avoiding the burden of biomarker testing, over a potentially higher risk of over- or undertreatment if relying only on symptoms, may consider choosing interval symptom-based treatment adjustment when being treated for active symptoms.

Implementation considerations:

- Interval biomarker assessment and treatment adjustment may be performed every 2–4 months in patients being treated for active symptoms.
- After resolution of symptoms (and normalization of biomarkers), endoscopic (and/or radiologic) evaluation should be performed to rule out active inflammation, typically 6–12 months after treatment initiation or adjustment. The patient may then transition to guidance for patients in symptomatic remission.

Summary of the Evidence

A biomarker-based evaluation strategy involves checking noninvasive biomarkers of inflammation in patients with symptomatically active CD to inform ongoing management; in contrast, symptom-based evaluation would involve treatment decisions being based solely on symptoms. One RCT, the CALM study, directly compared a biomarker-based evaluation strategy with symptom-based evaluation for patients with symptomatically active CD[14] (Table 8). In this multicenter, open-label RCT, Colombel et al[14] recruited adults with moderate to severely active nonstricturing, nonpenetrating CD, who were naïve to immune-suppressive therapy other than prednisone, and had endoscopic and biochemical evidence of inflammation. All patients were treated with prednisone and randomized after 9 weeks to "tight control," in which treatment escalation (initiation and subsequent escalation of adalimumab) was based on fecal calprotectin >250 μg/g and/or CRP >5 mg/L and/or symptoms suggestive of CD, vs "clinical management," in which treatment escalation was based on symptoms alone and was assessed every 12 weeks. Of the 244 included patients, more patients in the tight control group than in the clinical management group (37% vs 23%) achieved deep remission (defined as clinical remission [CDAI <150], endoscopic remission [CDEIS <4 and no deep ulcers], absence of draining fistula, and discontinuation of corticosteroids for 8 weeks or more) by 48 weeks.

Benefits and Harms (Downsides)

Symptom-based evaluation strategy. Potential benefit of a symptom-based monitoring strategy is the convenience of relying only on patient-reported outcomes, cost, and faster decision making. However, potential harms related to relying only on symptoms for treatment decisions are higher rates of inappropriate treatment adjustments or overtreatment and treatment-related complications; approximately 20%–35% of patients with gastrointestinal symptoms suggestive of CD may be in endoscopic remission.

Biomarker-based evaluation strategy. The potential benefits of a biomarker-based evaluation strategy include more accurate detection of inflammation than symptoms alone, to facilitate optimal treatment decisions and treatment escalation in patients with persistently elevated biomarkers, while simultaneously avoiding overtreatment. Potential harms of a biomarker-based evaluation strategy are the costs and inconvenience of sample collection, particularly stool-based tests, and potential delays in treatment that can happen due to the extra step of test completion and awaiting results.

Certainty of Evidence

From the CALM RCT, there was moderate certainty evidence supporting the use of biomarker-based evaluation strategy in patients with symptomatically active CD; evidence was rated down for imprecision due to low event rate.

Rationale

Using the GRADE Evidence-to-Decision framework, the guideline panel conditionally recommended in favor of a strategy that combines biomarkers and symptoms compared with symptom-based evaluation alone in patients with symptomatically active CD. The panel recognized that adding the extra step of biomarker testing in patients with symptomatically active CD may potentially delay treatment adjustments for patients, particularly those with limited access to health care resources. The panel recognized the value of shared decision making in these patients; some patients, particularly those with severe symptoms, who place high value on avoiding the burden of biomarker testing, may reasonably choose symptom-based evaluation for treatment decisions, acknowledging the potentially higher risk of inappropriate overtreatment with symptom-based evaluation alone. This may be particularly true if treatment decisions are considered low risk by the treating provider–patient team.

In the CALM RCT, interval biomarker assessment was performed every 3 months in patients with symptomatically active CD. The optimal management strategy in case of discrepancy between symptoms and biomarkers is unclear. In patients with typical symptoms suggestive of CD, normal biomarkers may not exclude lack of active inflammation, and endoscopic assessment may be a preferred approach. A subset of patients who were symptomatically active, and now have resolving symptoms on therapy, but have persistently elevated biomarkers, likely have ongoing inflammation. Treatment adjustments in response to elevated biomarkers are acceptable in this scenario. In this treat-to-target strategy in which symptoms and biomarkers normalize with iterative treatment adjustments in response to biomarker-based monitoring, endoscopic confirmation of remission is warranted to facilitate ongoing biomarker-based monitoring in asymptomatic patients, as recommended above.

Question 4: In patients with symptomatically active CD, at what fecal calprotectin, serum CRP, and Endoscopic Healing Index cutoffs can we accurately diagnose active inflammation, obviating routine endoscopic assessment?

Recommendation 6: In patients with CD with mild symptoms and elevated biomarkers of inflammation (fecal calprotectin >150 μg/g, CRP >5 mg/L), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, very low certainty in evidence*)

Implementation consideration:

- Lack of normalization of biomarkers (or persistently elevated biomarkers) in patients whose symptoms partially improve after initial treatment of active CD, likely suggests active inflammation, and may warrant treatment adjustment, without need for endoscopic (or radiologic) evaluation.

Recommendation 7: In patients with CD with mild symptoms and normal biomarkers of inflammation (fecal calprotectin <150 μg/g, CRP <5 mg/L), the AGA suggests endoscopic assessment of disease

Table 8. Grading of Recommendations Assessment, Development and Evaluation Evidence Profile for PICO (Patients, Intervention, Comparator, and Outcome) Question 3, Comparing Outcomes With Interval Biomarker-Based Monitoring vs Symptom-Based Monitoring to Improve Long-Term Outcomes in Patients With Symptomatically Active Crohn's Disease[a]

		Biomarker + symptom– vs symptom-based treatment adjustments in patients with symptomatic CD			
		Anticipated absolute effects (95% CI)			
Outcome No. of participants (studies)	Relative effect (95% CI)	Symptoms only	Symptoms + biomarker	Difference	Quality
Risk of achieving clinical and endoscopic remission at 12 mo 244 (1 RCT)	RR 1.61 (1.08–2.40)	Deep remission rate, % 23	37.0 (24.8–55.1)	14.0 more (1.8 more to 32.1 more)	⨁⨁⨁◯ MODERATE[b]

RR, relative risk.
[a]Patient or population: patients with symptomatically active CD. Setting: RCT. Intervention: treatment adjustment based on symptoms and/or biomarker elevation. Comparison: treatment adjustment based only on symptoms.
[b]Imprecision due to low event rate (73/244 achieved positive outcome; <200 events).

activity rather than empiric treatment adjustment. (*Conditional recommendation, very low certainty in evidence*)

Recommendation 8: In patients with CD with moderate to severe symptoms, the AGA suggests using fecal calprotectin >150 μg/g or CRP >5 mg/L to rule in active inflammation and inform treatment adjustment and avoid routine endoscopic assessment of disease activity. (*Conditional recommendation, low to moderate certainty in evidence*)

Recommendation 9: In patients with CD with moderate to severe symptoms with normal biomarkers of inflammation (fecal calprotectin <150 μg/g, CRP <5 mg/L), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, low certainty in evidence*)

Summary of the Evidence

Diagnostic performance of fecal calprotectin. Summary sensitivity and specificity of fecal calprotectin for detecting endoscopic inflammation is reported in Question 2.

High pretest probability scenario (patients with moderate to severe symptoms suggestive of CD flare [PRO2 >13 or PRO3 >21], with 80% prevalence of inflammation). In applying these fecal calprotectin cutoffs in high pretest probability scenarios, approximately 6.6%, 5.6%, and 5.2% patients (FP rate) with fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g, respectively, may be misclassified as having endoscopic activity when actually these patients are in endoscopic remission (Table 9). In contrast, fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g in this high pretest probability scenario, had significantly high rates of being FN (9.6%, 15.2%, and 19.2%, respectively), that is, a significant proportion of symptomatic patients with fecal calprotectin below thresholds who have endoscopic activity, may be incorrectly classified as being in endoscopic remission.

Intermediate pretest probability scenario (patients with mild symptoms of CD [PRO2 score 8–13 or PRO3 score 13–21], with 65% prevalence of endoscopic inflammation). In an intermediate pretest probability scenario, approximately 11.5%, 9.8%, and 9.1% of patients (FP rate) with fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g, respectively, may be misclassified as having endoscopic activity when they are actually in endoscopic remission (Table 9). In addition, fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g in this intermediate pretest probability scenario had high rates of being FN (7.8%, 12.3%, and 15.6%, respectively), that is, a substantial proportion of mildly symptomatic patients who have endoscopic activity may be incorrectly classified as being in endoscopic remission.

Diagnostic performance of serum C-reactive protein. Summary sensitivity and specificity of serum CRP for detecting active endoscopic inflammation is reported in Question 2.

High pretest probability scenario (patients with moderate to severe symptoms suggestive of CD flare [PRO2 >13 or PRO3 >21], with 80% prevalence of inflammation). In applying this cutoff (elevated CRP, generally >5 mg/L) to a high pretest probability scenario, only approximately 5.4% patients (FP rate) with elevated CRP may be misclassified as having endoscopic activity while in endoscopic remission (Table 10). In contrast, normal CRP (<5 mg/L) had significantly high rates of being FN (26.4%), that is, a high proportion of symptomatic patients with normal CRP who have endoscopic activity may be incorrectly classified as being in endoscopic remission.

Intermediate pretest probability scenario (patients with mild symptoms of CD [PRO2 score 8–13 or PRO3 score 13–21], with 65% prevalence of endoscopic inflammation). In an intermediate pretest probability scenario, FP

Table 9. Grading of Recommendations Assessment, Development and Evaluation Evidence Profile for PICO (Patients, Intervention, Comparator, and Outcome) Question 4, Comparing Cutoffs for Fecal Calprotectin in Patients With Symptomatically Active Crohn's Disease[a]

	No. of results per 1000 patients tested (95% CI)						
	Mild symptoms (prevalence 65%)			Moderate to severe symptoms (prevalence 80%)			
Test result	fCal >50 μg/g	fCal >150 μg/g	fCal>250 μg/g	fCal >50 μg/g	fCal >150 μg/g	fCal >250 μg/g	Comments
TPs (patients correctly diagnosed as having endoscopic activity)	572 (514–611)	527 (481–566)	494 (455–533)	704 (632–752)	648 (592–696)	608 (560–656)	TPs would be eligible to undergo treatment adjustment, which may decrease disease-related complications and morbidity, without being subject to risks and invasive testing with endoscopy.
FNs (patients incorrectly labeled as being in endoscopic remission, when actually they have endoscopic activity)	78 (39–136)	123 (84–169)	156 (117–195)	96 (48–168)	152 (104–208)	192 (144–240)	FNs may be falsely reassured, undertreated, or mistreated (as not having CD flare), potentially leading to increased disease-related complications and morbidity.
GRADE certainty of evidence	LOW[b,c]	VERY LOW[b,d]	VERY LOW[b,d]	LOW [b,c]	VERY LOW[b,d]	VERY LOW[b,d]	
TNs (patients correctly diagnosed as being in endoscopic remission)	235 (179–280)	252 (214–284)	259 (235–280)	134 (102–160)	144 (122–162)	148 (134–160)	TNs would be reassured and obviate the need for invasive testing with endoscopy, although they may need to undergo serial assessment of biomarker at periodic intervals.
FPs (patients incorrectly labeled as having endoscopic activity)	115 (70–171)	98 (66–136)	91 (70–115)	66 (40–98)	56 (38–78)	52 (40–66)	FPs may undergo unnecessary treatment adjustment and have treatment-related complications.
GRADE certainty of evidence	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	LOW[b,c]	LOW[b,c]	LOW[b,c]	

fCal, fecal calprotectin.
[a]Population/setting: adults with CD with symptoms suggestive of active CD, mild symptoms (PRO2 8–13, or PRO3 13–21) with observed prevalence of endoscopically active disease of 65%; moderate to severe symptoms (PRO2 >13 or PRO3 >21) with observed prevalence of endoscopically active disease of 80%. Pooled sensitivity/specificity fCal with cutoff <50 μg/g: sensitivity, 88% (95% CI, 79%–94%); specificity, 67% (95% CI, 51%–80%), 16 studies. Pooled sensitivity/specificity fCal with cutoff <150 μg/g: sensitivity, 81% (95% CI, 74%-87%); specificity, 72% (95% CI, 61%–81%), 13 studies. Pooled sensitivity/specificity fCal with cutoff <250 μg/g: sensitivity, 76% (9% CI, 70%–82%); specificity, 74% (95% CI, 67%–80%), 14 studies. Reference test: colonoscopy.
[b]High unexplained heterogeneity, selective inclusion of studies reporting cutoffs.
[c]Serious imprecision because 95% CI crosses maximal tolerable FN threshold of <5%.
[d]Very serious imprecision because point estimate and 95% CI are higher than maximal tolerable FN threshold.
[e]Very serious imprecision because point estimate is higher than maximal tolerable FP threshold.

Table 10. Grading of Recommendations Assessment, Development and Evaluation Evidence Profile for PICO (Patients, Intervention, Comparator, and Outcome) Question 4, Comparing Cutoffs for Serum C-Reactive Protein in Patients With Symptomatically Active Crohn's Disease[a]

	No. of results per 1000 patients tested (95% CI)		
	Mild symptoms (prevalence 65%)	Moderate-severe symptoms (prevalence 80%)	
Test result	Elevated CRP	Elevated CRP	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	436 (351–501)	536 (432–616)	TPs would be eligible to undergo treatment adjustment, which may decrease disease-related complications and morbidity, without being subject to risks and invasive testing with endoscopy.
FNs (patients incorrectly labeled as being in endoscopic remission or having only mild endoscopic activity, when actually they have moderate to severe endoscopic activity)	214 (149–299)	264 (184–368)	FNs may be falsely reassured, undertreated, or mistreated (as not having CD flare), potentially leading to increased disease-related complications and morbidity.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	
TNs (patients correctly diagnosed as being in endoscopic remission or having only mild endoscopic activity)	256 (227–280)	146 (130–160)	TNs would be reassured and obviate the need for invasive testing with endoscopy, although they may need to undergo serial assessment of biomarker at periodic intervals.
FPs (patients incorrectly labeled as having moderate to severe endoscopic activity, when actually they are in endoscopic remission or have only mild endoscopic activity)	94 (70–123)	54 (40–70)	FPs may undergo unnecessary treatment adjustment and have treatment-related complications.
GRADE certainty of evidence	VERY LOW[b,d]	LOW[b,e]	—

[a]Population/setting: adults with CD with symptoms suggestive of active CD, mild symptoms (PRO2 8–13, or PRO3 13–21) with observed prevalence of endoscopically active disease of 65%; moderate to severe symptoms (PRO2 >13 or PRO3 >21) with observed prevalence of endoscopically active disease of 80%. Pooled sensitivity of CRP <5 mg/L, 66.7% (95% CI, 54.4%–77.1%), 20 studies. Pooled specificity of CRP <5 mg/L, 73.1% (95% CI, 64.7%–80.1%), 20 studies. Reference test: colonoscopy.
[b]Serious inconsistency due to high unexplained for summary sensitivity/specificity, inconsistent reporting by location.
[c]Very serious imprecision because point estimate is higher than maximal tolerable FN threshold.
[d]Very serious imprecision because point estimate is higher than maximal tolerable FP threshold.
[e]Serious imprecision because 95% CI crosses maximal tolerable FP threshold of <5%.

rate of elevated CRP was 9.4% and FN rate of normal CRP was 21.4% (Table 10), that is, a high proportion of symptomatic patients who have endoscopic activity may be incorrectly classified as being in endoscopic remission.

Certainty of Evidence

Even though there were no direct data comparing how different biomarker cutoffs and accompanying treatment decisions impact downstream patient-important outcomes, we did not rate down for indirectness because the presence of endoscopic activity is a close surrogate for unfavorable patient outcomes and an indication for treatment adjustment.

Fecal calprotectin. There was low certainty of evidence supporting the use of any proposed fecal calprotectin cutoff to rule in endoscopic inflammation in a high pretest probability setting (evidence rated down for inconsistency due to selective inclusion of studies reporting specific cutoffs and high heterogeneity for summary sensitivity and specificity and imprecision because 95% CI exceeded the maximal

tolerable FP rate of 5%). In contrast, in the intermediate probability scenario of patients with mild symptoms, there was very low certainty of evidence supporting the use of any proposed fecal calprotectin cutoff to rule in endoscopic inflammation due to unacceptably high rates of FP (inconsistency and very serious imprecision because both the point estimate and 95% CI are higher than FP threshold of 5%). Similarly, in both the high and intermediate probability scenarios, there was very low certainty of evidence supporting the use of fecal calprotectin <150 μg/g or <250 μg/g as cutoffs to rule out endoscopic inflammation, due to unacceptably high rates of FN (very serious imprecision) and inconsistency. However, the fecal calprotectin cutoff of <50 μg/g performed slightly better, although it still was rated low certainty of evidence (inconsistency and imprecision because 95% CI exceeded the maximal tolerable FN rate of 5%).

Serum C-reactive protein. There was low certainty of evidence supporting the use of elevated CRP to rule in endoscopic inflammation in the high pretest probability setting. Evidence was rated down for inconsistency due to selective inclusion of studies reporting specific cutoffs and high heterogeneity for summary sensitivity and specificity, and imprecision because 95% CI exceeded the maximal tolerable FP rate of 5%. In an intermediate pretest probability scenario, the certainty of evidence was very low for ruling out endoscopic inflammation (inconsistency and very serious imprecision because both the point estimate and 95% CI were higher than FP threshold of 5%). In contrast, in both the intermediate and high probability scenario, there was very low certainty of evidence supporting the use of normal serum CRP to rule out endoscopic inflammation (inconsistency and very serious imprecision).

Rationale

The guideline panel determined *a priori* the maximal tolerable FP thresholds at 5% for patients with symptomatically active CD. However, the panel deemed that there may be circumstances when patients and providers may be willing to accept higher rates of FP, depending on risk of downstream consequences, including the nature of treatment adjustment. Thus, the panel promotes shared decision making with conditional recommendations. For example, while endoscopic evaluation would be warranted for most patients with mild symptoms, decisions based on elevated biomarkers may be acceptable if there are logistical delays in obtaining endoscopic evaluation and if patients and providers are willing to accept negative consequences associated with low-risk treatment adjustments.

As noted earlier, for ease of implementation in clinical practice, the guideline panel felt that choosing a single fecal calprotectin cutoff (>150 μg/g) that is broadly applicable across a wide range of clinical scenarios is preferable, rather than reporting different cutoffs for different scenarios. Higher fecal calprotectin cutoffs may have modestly lower rates of FP with modest improvement in confidence of decision making.

Postoperative Management of Crohn's Disease

Question 5: In patients with CD in surgically induced remission, at what fecal calprotectin, serum CRP, and Endoscopic Healing Index cutoffs can we accurately rule out postoperative endoscopic recurrence, obviating routine endoscopic assessment?

Recommendation 10: In asymptomatic patients with CD after surgically induced remission within the past 12 months, who are at low risk of postoperative recurrence or who have 1 or more risk factors for recurrence but are on postoperative pharmacologic prophylaxis, the AGA suggests using fecal calprotectin <50 μg/g to avoid routine endoscopic assessment of disease activity. (*Conditional recommendation, moderate certainty in evidence*)

Comment: Patients, particularly those with multiple prior surgeries, and/or with failure of multiple advanced therapies before surgery, who value more accurate assessment of endoscopic recurrence over the inconvenience and costs of colonoscopy, may reasonably choose endoscopic assessment of disease activity within 12 months after surgery.

Recommendation 11: In asymptomatic patients with CD after surgically induced remission within the past 12 months, who are at high baseline risk of recurrence and are not receiving postoperative pharmacologic prophylaxis, the AGA suggests endoscopic evaluation, rather than relying solely on biomarkers, for assessing endoscopic recurrence. (*Conditional recommendation, low to moderate certainty in evidence*)

Implementation considerations:

- Risk stratification schemes to classify patient's risk of endoscopic recurrence after surgically induced remission are not well-defined. Risk factors typically associated with low risk of recurrence 6–12 months after surgically induced remission include older age at surgery (older than 50 years), nonsmoking, longstanding disease (more than 10 years), and first surgery for a short segment of fibrostenotic disease ($<$10–20 cm).[31–33] Risk factors typically associated with high risk of recurrence 6–12 months after surgically induced remission include 2 or more prior surgeries, penetrating or perianal disease, smoking, young age at surgery, with long segment of small bowel resection; however, these risk factors may not be additive.
- In patients at low baseline risk of recurrence, who are also receiving postoperative pharmacologic prophylaxis, fecal calprotectin <150 μg/g may also rule out endoscopic recurrence.
- Normal CRP in patients with asymptomatic CD in surgically induced remission is not able to rule out endoscopic recurrence accurately.

- There are limited data on ongoing biomarker monitoring alone in patients with CD in surgically induced remission. Colonoscopic evaluation may be warranted beyond 12 months after surgery in patients when biomarker-based monitoring is being pursued.

Summary of the Evidence

Diagnostic performance of fecal calprotectin. We conducted a systematic review to identify cross-sectional and cohort studies in patients with established CD in surgically induced remission reported the diagnostic accuracy of fecal calprotectin for detecting endoscopic recurrence with active inflammation reported as Rutgeerts score i2 or higher in most studies. To minimize bias due to selective reporting of optimized cutoffs, we included only studies that reported diagnostic accuracy of preselected fecal calprotectin cutoffs or reported the performance across 2 or more predetermined cutoffs. Overall, 22 cohorts met these criteria. Using this approach, the sensitivity and specificity of fecal calprotectin cutoff of 50 ± 50 μg/g were 86% (95% CI, 75%–93%) and 50% (95% CI, 34%–66%), respectively, based on 11 cohorts; corresponding sensitivity and specificity of 150 ± 50 μg/g cutoff (6 cohorts) were 64% (95% CI, 51%–75%) and 71% (95% CI, 63%–77%), respectively, and of 250 ± 50 μg/g cutoff (6 cohorts) were 52% (95% CI, 40%–64%) and 79% (95% CI, 66%–88%), respectively (Supplementary Figure 7).

Low pretest probability scenario (low baseline risk of endoscopic recurrence, on pharmacologic prophylaxis after surgery, estimated 10% prevalence of endoscopic recurrence). In applying fecal calprotectin diagnostic cutoffs to this low pretest probability scenario, approximately 1.4%, 3.6%, and 4.8% patients (FN rate) with fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g, respectively, may be misclassified as being in endoscopic remission (Rutgeerts score i0 or i1) when they actually have endoscopic recurrence (Rutgeerts score i2 or higher) (Table 11). In contrast, elevated fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g in this low pretest probability scenario had significantly high rates of being FP (45.0%, 26.1%, and 18.9%, respectively), that is, a significant proportion of patients who are in endoscopic remission may be incorrectly classified as having endoscopic recurrence (Supplementary Figure 8).

Intermediate pretest probability scenario (asymptomatic patients with high baseline risk of endoscopic recurrence who are receiving postoperative prophylaxis or patients with low baseline risk of endoscopic recurrence who are not receiving pharmacologic prophylaxis, estimated 30% prevalence of endoscopic recurrence). In this intermediate pretest probability scenario, approximately 4.2%, 10.8%, and 14.4% patients (FN rate) with fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g, respectively, may be misclassified as being in endoscopic remission when they actually have endoscopic recurrence (Table 11). In contrast, elevated fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g in this intermediate pretest probability scenario, had significantly high rates of being FP (35.0%, 20.3%, and 14.7%, respectively), that is, a significant proportion of patients who are in endoscopic remission may be incorrectly classified as having endoscopic recurrence.

High pretest probability scenario (asymptomatic patients with high baseline risk of endoscopic recurrence and are not receiving postoperative pharmacologic prophylaxis, with estimated 60% prevalence of endoscopic recurrence). In this high pretest probability scenario, approximately 8.4%, 21.6%, and 28.8% of patients (FN rate) with fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g, respectively, may be misclassified as being in endoscopic remission when they actually have endoscopic recurrence (Table 11). In contrast, elevated fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g in this high pretest probability scenario had significantly high rates of being FP (20.0%, 11.6%, and 8.4%, respectively), that is, a significant proportion of patients who are in endoscopic remission may be incorrectly classified as having endoscopic recurrence.

Diagnostic performance of serum C-reactive protein. We only relied on studies that simultaneously reported both fecal calprotectin and CRP data. We identified 4 studies reporting the diagnostic accuracy of serum CRP for detecting postoperative endoscopic recurrence. Most studies reported endoscopic recurrence as Rutgeerts score i2 or higher. Summary sensitivity and specificity of elevated CRP for detecting endoscopic recurrence were 30% (95% CI, 21%–40%) and 90% (95% CI, 84%–94%).

Low pretest probability scenario (patients with low baseline risk who are receiving pharmacologic prophylaxis, estimated 10% prevalence of endoscopic recurrence). In applying this cutoff (elevated CRP, generally >5 mg/L) to a low pretest probability scenario, approximately 7.0% patients (FN rate) with normal CRP (<5 mg/L) may be misclassified as having endoscopic remission when they actually have endoscopic recurrence (Table 12). In contrast, elevated CRP (>5 mg/L) in this low pretest probability scenario had moderate rates of being FP (9.0%), that is, 9.0% patients who have endoscopic remission may be incorrectly classified as having endoscopic recurrence.

Intermediate pretest probability scenario (asymptomatic patients with high baseline risk of recurrence, receiving postoperative prophylaxis or low baseline risk patients not receiving prophylaxis, estimated 30% prevalence of endoscopic recurrence). In an intermediate pretest probability scenario, approximately 21.0% of patients (FN rate) with normal CRP (<5 mg/L) may be misclassified as having endoscopic remission when they actually have endoscopic recurrence (Rutgeerts score i2 or higher) (Table 12). In contrast, elevated CRP (>5 mg/L), in this intermediate pretest probability scenario, had moderate rates of being FP (7.0%).

High pretest probability scenario (asymptomatic patients with high risk of recurrence not receiving postoperative prophylaxis, with estimated 60% prevalence of endoscopic recurrence). In a high pretest probability scenario, approximately 42.0% of patients (FN rate) with normal CRP (<5 mg/L) may be misclassified as having endoscopic remission when they actually have endoscopic

Table 11. Grading of Recommendations Assessment, Development and Evaluation Evidence Profile for PICO (Patients, Intervention, Comparator, and Outcome) Question 5, Comparing Cutoffs for Fecal Calprotectin in Asymptomatic Patients With CD Being Followed After Surgically Induced Remission[a]

	No. of results per 1000 patients tested (95% CI)						
	Low pretest probability (prevalence 10%)			Intermediate pretest probability (prevalence 30%)			
Test result	fCal <50 μg/g	fCal <150 μg/g	fCal <250 μg/g	fCal <50 μg/g	fCal <150 μg/g	fCal <250 μg/g	Comments
TPs (patients correctly diagnosed as having endoscopic activity)	86 (75–93)	64 (51–75)	52 (40–64)	258 (225–279)	192 (153–225)	156 (120–192)	TPs would be eligible to undergo treatment adjustment, which may decrease disease-related complications and morbidity, without being subject to risks and invasive testing with endoscopy.
FNs (patients incorrectly labeled as being in endoscopic remission, when actually they have endoscopic activity)	14 (7–25)	36 (25–49)	48 (36–60)	42 (21–75)	108 (75–147)	144 (108–180)	FNs would be falsely reassured and may be at higher risk of disease complications/ progression and flare due to undertreatment and missed opportunity to treat.
GRADE certainty of evidence	MODERATE[b]	MODERATE[b]	LOW[b,c]	LOW[b,c]	VERY LOW[b,d]	VERY LOW[b,d]	
TNs (patients correctly diagnosed as being in endoscopic remission)	450 (306–594)	639 (567–693)	711 (594–792)	350 (238–462)	497 (441–539)	553 (462–616)	TNs would be reassured and obviate the need for invasive testing with endoscopy, although they may need to undergo serial assessment of biomarker at periodic intervals.
FPs (patients incorrectly labeled as having endoscopic activity, when actually they are in endoscopic remission)	450 (306–594)	261 (207–333)	189 (108–306)	350 (238–462)	203 (161–259)	147 (84–238)	FPs may receive unnecessary testing (endoscopy) and/or treatment adjustment, and have avoidable anxiety, potential testing- or treatment-related complications, and excessive resource utilization.
GRADE certainty of evidence	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	—

Table 11. Continued

	No. of results per 1000 patients tested (95% CI)			
	High pretest probability (prevalence 60%)			
	fCal <50 μg/g	fCal <150 μg/g	fCal <250 μg/g	
FNs (patients incorrectly labeled as being in endoscopic remission, when actually they have endoscopic activity)	84 (42–150)	216 (150–294)	288 (216–360)	FNs would be falsely reassured and may be at higher risk of disease complications/ progression and flare due to undertreatment and missed opportunity to treat.
GRADE certainty of evidence	LOW[b,c]	VERY LOW[b,d]	VERY LOW[b,d]	
FPs (patients incorrectly labeled as having endoscopic activity, when actually they are in endoscopic remission)	200 (136–264)	116 (92–148)	84 (48–136)	FPs may receive unnecessary testing (endoscopy) and/or treatment adjustment, and have avoidable anxiety, potential testing- or treatment-related complications, and excessive resource utilization.
GRADE certainty of evidence	VERY LOW[b,e]	VERY LOW[b,e]	VERY LOW[b,e]	—

[a]Population/setting: patients with CD in surgically induced remission, low pretest probability/likelihood of having endoscopic recurrence such as patients at low baseline risk of recurrence (no high-risk features) and on postoperative prophylaxis, 10%; intermediate pretest probability/likelihood of having endoscopic recurrence such as patients at low baseline risk of recurrence (no high-risk features), who are not on postoperative prophylaxis or patients at high baseline risk of recurrence who are receiving postoperative prophylaxis, 30%; high pretest probability/likelihood of having endoscopic recurrence such as patients at high baseline risk of recurrence who are not receiving postoperative prophylaxis, 60%. Pooled sensitivity/specificity fCal with cutoff <50 μg/g: sensitivity, 86% (95% CI, 75%–93%); specificity, 50% (95% CI, 34%–66%), 11 studies. Pooled sensitivity/specificity fCal with cutoff <150 μg/g: sensitivity, 64% (95% CI, 51%–75%); specificity, 71% (95% CI, 63%–77%), 6 studies. Pooled sensitivity/specificity fCal with cutoff <250 μg/g: sensitivity, 52% (95% CI, 40%–64%); specificity, 79% (95% CI, 66%–88%), 6 studies. Reference test: colonoscopy.
[b]High unexplained heterogeneity, selective inclusion of studies reporting cutoffs.
[c]Serious imprecision because 95% CI crosses maximal tolerable FN threshold of <5%.
[d]Very serious imprecision because point estimate and 95% CI are higher than maximal tolerable FN threshold.
[e]Very serious imprecision because point estimate is higher than maximal tolerable FP threshold.

Table 12. Grading of Recommendations Assessment, Development and Evaluation Evidence Profile for PICO (Patients, Intervention, Comparator, and Outcome) Question 5, Comparing Cutoffs for (Serum C-Reactive Protein in Asymptomatic Patients With Crohn's Disease Being Followed After Surgically Induced Remission[a]

	No. of results per 1000 patients tested (95% CI)			
	Normal CRP			
Test result	Low pretest probability (10%)	Intermediate pretest probability (30%)	High pretest probability (60%)	Comments
FNs (patients incorrectly labeled as being in endoscopic remission, when actually they have endoscopic activity)	70 (60 to 79)	210 (180 to 237)	420 (360 to 474)	FNs would be falsely reassured and may be at higher risk of disease complications/progression and flare due to undertreatment and missed opportunity to treat.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	VERY LOW[b,c]	
FPs (patients incorrectly labeled as having endoscopic activity, when actually they are in endoscopic remission)	90 (54 to 144)	70 (42 to 112)	40 (24 to 64)	FPs may receive unnecessary testing (endoscopy) and/or treatment adjustment, and have avoidable anxiety, potential testing- or treatment-related complications, and excessive resource utilization.
GRADE certainty of evidence	VERY LOW[b,d]	LOW[b,e]	LOW[b,e]	

Pooled sensitivity/specificity fecal calprotectin with cutoff <150 μg/g: sensitivity, 64% (95% CI, 51%–75%); specificity, 71% (95% CI, 63%–77%), 6 studies. Pooled sensitivity/specificity fecal calprotectin with cutoff <250 μg/g: sensitivity, 52% (95% CI, 40%–64%); specificity, 79% (66%–88%), 6 studies. Reference test: colonoscopy.
[a]Population/setting: patients with CD in surgically induced remission, low pretest probability/likelihood of having endoscopic recurrence, such as patients at low baseline risk of recurrence (no high-risk features) and on postoperative prophylaxis, 10%; intermediate pretest probability/likelihood of having endoscopic recurrence, such as patients at low baseline risk of recurrence (no high-risk features), who are not on postoperative prophylaxis or patients at high baseline risk of recurrence who are receiving postoperative prophylaxis, 30%; high pretest probability/likelihood of having endoscopic recurrence, such as patients at high baseline risk of recurrence who are not receiving postoperative prophylaxis, 60%. Pooled sensitivity/specificity fecal calprotectin with cutoff <50 μg/g: sensitivity, 86% (95% CI, 75%–93%); specificity, 50% (95% CI, 34%–66%), 11 studies.
[b]High unexplained heterogeneity, selective inclusion of studies reporting cutoffs.
[c]Very serious imprecision because point estimate is higher than maximal tolerable FN threshold.
[d]Very serious imprecision because point estimate is higher than maximal tolerable FP threshold.
[e]Serious imprecision because 95% CI crosses maximal tolerable FP threshold of <5%.

recurrence (Table 12). In contrast, elevated CRP (>5 mg/L), in this high pretest probability scenario, had low rates of being FP (4.0%).

Certainty of the Evidence

Fecal calprotectin. There was moderate certainty of evidence supporting the use of fecal calprotectin <50 μg/g and <150 μg/g to rule out postoperative recurrence in a low pretest probability scenario (evidence rated down for inconsistency due to selective inclusion of studies reporting specific cutoffs and high heterogeneity for summary sensitivity and specificity) and low certainty of evidence supporting the use of fecal calprotectin <250 μg/g to rule out postoperative recurrence in a low pretest probability scenario (evidence rated down for inconsistency and imprecision because 95% CI of the FN crosses the established threshold of 5%). In contrast, in the intermediate and high pretest probability scenarios, low certainty of evidence supported the use of fecal calprotectin <50 μg/g (inconsistency and imprecision), and very low certainty of evidence supported the use of fecal calprotectin <150 μg/g and <250 μg/g to rule out postoperative recurrence (inconsistency and very serious imprecision because both the point estimate and 95% CI are higher than FN threshold of 5%).

In all probability scenarios, there was very low certainty of evidence supporting the use of fecal calprotectin >50 μg/g, >150 μg/g, or >250 μg/ to rule in endoscopic recurrence in asymptomatic patients due to unacceptably high rates of FP (very serious imprecision) and inconsistency. However, the fecal calprotectin cutoff of >250 μg/g performed slightly better, although it was still rated low certainty of evidence (inconsistency and imprecision because 95% CI exceeded the maximal tolerable FP rate of 5%).

Serum C-reactive protein. In all probability scenarios, there was very low certainty of evidence supporting the use of normal CRP <5 mg/L to rule out endoscopic recurrence in asymptomatic patients due to unacceptably high rates of FN (very serious imprecision) and

inconsistency. In contrast, low certainty supported the use of elevated CRP >5 mg/L to rule in endoscopic recurrence in asymptomatic patients in a high and intermediate pretest probability scenario (inconsistency and imprecision because 95% CI exceeded the maximal tolerable FP rate of 5%); there was very low certainty evidence supporting its use in a low pretest probability scenario.

Rationale

Using the GRADE Evidence-to-Decision framework, the guideline panel conditionally recommended in favor of a strategy that uses fecal calprotectin–based monitoring, using a cutoff of <50 μg/g, over routine endoscopic evaluation to rule out postoperative recurrence in asymptomatic patients with CD after surgically induced remission within the past 12 months and who are either receiving postoperative pharmacologic prophylaxis or at low baseline risk of postoperative recurrence (regardless of postoperative prophylaxis). In asymptomatic patients with CD after surgically induced remission within the past 12 months, who are at high baseline risk of recurrence, and are not receiving postoperative pharmacologic prophylaxis, the panel conditionally recommended endoscopic evaluation rather than relying solely on biomarkers for assessing endoscopic recurrence.

The guideline panel recognizes the challenge in clinical practice of using a different fecal calprotectin cutoff (<50 μg/g) in the evaluation of asymptomatic patients with CD after surgically induced remission within the past 12 months. It is worth noting that fecal calprotectin performed well at <50 μg/g to rule out endoscopic recurrence and active inflammation in patients at low and intermediate baseline risk of recurrence. In patients at low baseline risk of recurrence, who are also receiving postoperative pharmacologic prophylaxis, fecal calprotectin <150 μg/g may also rule out endoscopic recurrence. However, in patients at high baseline risk of recurrence, normal fecal calprotectin does not rule out endoscopic recurrence and an endoscopic evaluation is recommended in these patients over biomarker-based assessment.

In contrast to the utility of a normal fecal calprotectin value, the test performance of elevated fecal calprotectin was not sufficient to recommend relying only on this to make treatment decisions in asymptomatic patients with CD after surgically induced remission within the past 12 months, regardless of pretest probability scenarios. In this setting, the guideline panel recommends endoscopic evaluation to confirm the presence and severity of endoscopic recurrence before treatment adjustments. The data for CRP did not support its use (either normal or elevated value) to determine endoscopic recurrence accurately in patients with asymptomatic CD in surgically induced remission.

There are limited data on ongoing biomarker monitoring alone in patients with CD in surgically induced remission. The panel noted that, given the limited data, endoscopic evaluation may be warranted beyond 12 months after surgery in patients where biomarker-based monitoring is being pursued. For patients with surgically induced remission with symptoms, the panel recommends management strategy outlined in Recommendations 6–9 for patients with mild or moderate to severe symptoms.

Endoscopic Healing Index (Monitr)

Question 6: In patients with CD in symptomatic remission, at what EHI cutoff can we accurately rule out active inflammation, obviating routine endoscopic assessment?

Question 7: In patients with symptomatically active CD, at what EHI cutoff can we accurately diagnose active inflammation, obviating routine endoscopic assessment?

Question 8: In patients with CD in surgically induced remission, at what EHI cutoff can we accurately rule out postoperative endoscopic recurrence, obviating routine endoscopic assessment?

Recommendation 12: In patients with CD, the AGA suggests neither in favor of nor against the use of EHI (Monitr) for monitoring inflammation and treatment decisions. (*No recommendation, Knowledge gap*)

Summary of the Evidence

Diagnostic performance of Endoscopic Healing Index in luminal Crohn's disease. The EHI (Monitr) measures 13 proteins in blood (ie, ANG1, ANG2, CRP, SAA1, IL7, EMMPRIN, MMP1, MMP2, MMP3, MMP9, TGFA, CEACAM1, and VCAM1) and was developed as a diagnostic test to reflect the severity of endoscopic inflammation in CD. A single derivation–validation study was identified that evaluated the test performance of EHI in 2 validation cohorts.[34] The first validation cohort, TAILORIX (Tailored Treatment With Infliximab for Active Crohn's Disease), consisted of 116 patients with prospectively collected data, 26% of whom were in endoscopic remission; 10% had a history of IBD surgery. The second validation cohort included samples prospectively collected from a tertiary referral center (University of California San Diego); 46% had a history of IBD surgery.

Two cutoffs were determined for evaluation of luminal CD. A cutoff of EHI <20 was optimized to rule out active inflammation, defined as SES-CD <3. In the TAILORIX cohort, the sensitivity and specificity of EHI <20 were 96% and 64%, respectively. In the University of California San Diego cohort, the sensitivity and specificity of EHI <20 were 92% and 42%, respectively. A cutoff of EHI >50 was optimized to rule in active inflammation, defined as SES-CD ≥3. In the TAILORIX cohort, the sensitivity and specificity of EHI >50 were 36% and 100%, respectively. In the University of California San Diego cohort, the sensitivity and specificity of EHI >50 were 35% and 91%, respectively.

With this optimized test performance, EHI <20 had very low rates of FN in asymptomatic patients with luminal CD,

regardless of whether patients were in known (FN 1.2%) or unknown endoscopic remission status (FN 2.7%). In these clinical scenarios, EHI >20 had very high rates of FP (>25%), implying that a high proportion of asymptomatic patients in endoscopic remission will be incorrectly classified as having endoscopic activity when EHI >20 (Table 13). In patients with symptomatically active CD, with mild symptoms or with moderate to severe symptoms, EHI >50 had very low rates of FP (<2%). In these symptomatic patients, EHI <50 had very high rates of FN (>40%), implying that a high proportion of symptomatic patients with endoscopically active disease will be incorrectly classified as being in remission when EHI <50 is used as a cutoff (Table 14). Although the test cutoffs were optimized, a large number of results of the test were in the 20–50 range, making them indeterminate.

Diagnostic performance of Endoscopic Healing Index in postoperative Crohn's disease. A single study was identified evaluating the use of EHI in postoperative CD.[35] This study was a secondary analysis of the POCER (Postoperative Crohn's Endoscopic Recurrence) trial. At 6 months, the sensitivity and specificity of EHI <20 for endoscopic recurrence (Rutgeerts score i2 or higher) were 82% and 50%, respectively (Table 15). At this optimized cutoff, EHI <20 was able to accurately rule out endoscopic recurrence (FN rate <6%) in patients at low- and intermediate pretest probability scenarios, including asymptomatic patients with CD after surgically induced remission within the past 12 months and who are either receiving postoperative pharmacologic prophylaxis or at low baseline risk of postoperative recurrence (regardless of postoperative prophylaxis).

Certainty of the Evidence

The overall body of evidence supporting the use of EHI in different clinical scenarios was rated as very low quality due to overall paucity of studies. Only 1 derivation and 2 accompanying validation studies examined the performance of EHI in patients with luminal CD; there was only 1 study evaluating the performance of EHI in patients with postoperative CD. Consequently, the body of evidence was rated down for very serious imprecision and possible reporting or publication bias.

Rationale

Using the GRADE Evidence-to-Decision framework, the guideline panel decided to make a recommendation neither in favor of nor against the use of the EHI test in CD. In arriving at this recommendation, the guideline panel weighed the performance of the test in the reported validation cohorts against the paucity of independent data, despite the test being commercially available since 2020. Based on this, together with limited access and feasibility and risk of exacerbating inequity for this proprietary test, the guideline panel opted not to make a recommendation in favor of, or against, its use, identifying this as a knowledge gap. The availability of more generalizable data demonstrating high accuracy from independent data sets in sufficiently heterogenous populations, as well as evidence of its feasibility and cost-effectiveness relative to other widely available tests, would merit reconsideration of the recommendation.

Biomarker- vs Endoscopy-Based Monitoring Strategy in Crohn's Disease

Question 9: In patients with established CD, is interval biomarker-based monitoring superior to endoscopy-based monitoring to improve long-term outcomes?

Recommendation 13: In patients with CD, the AGA makes no recommendation in favor of, or against, a biomarker-based monitoring strategy over an endoscopy-based monitoring strategy to improve long-term outcomes. (*No recommendation, Knowledge gap*)

Summary of the Evidence

A biomarker-based monitoring strategy involves routine assessment of symptoms and noninvasive biomarkers of inflammation in patients with CD in symptomatic remission to inform ongoing management. In this situation, normalization of biomarkers is an adequate treatment target—asymptomatic patients with normal biomarkers would continue current management without endoscopy, whereas those with elevated biomarkers would undergo endoscopy. In contrast, an endoscopy-based monitoring strategy involves routine endoscopic assessment to confirm endoscopic remission of CD periodically. Supplementary Figure 9 lays out the schematic for proposed comparison. We did not identify any RCTs that compared a biomarker-based monitoring strategy with an endoscopy-based monitoring strategy. Normalization of CRP and reduction of fecal calprotectin are recognized as short-term treatment targets in managing CD in expert consensus statements, assessed early in treatment course. Early achievement of these biomarker outcomes is associated with favorable longer-term outcomes, including risk of relapse as well as likelihood of achieving endoscopic improvement. Potential benefits of a biomarker-based monitoring strategy are convenience and low resource utilization due to avoidance of routine and recurrent endoscopic assessment. Potential harms of a biomarker-based monitoring strategy are insufficient assessment and suboptimal performance for achieving deeper remission end points, such as complete endoscopic or transmural remission, which may be associated with more favorable long-term outcomes. Hence, the guideline panel felt there was insufficient evidence to inform between the choice of a biomarker-based monitoring strategy and an endoscopy-based monitoring strategy in patients with CD in symptomatic remission. This was identified as a knowledge gap that warrants further study.

Limitations of Current Evidence and Future Directions

The evidence panel identified numerous knowledge gaps in the literature where there was insufficient data to inform recommendations.

Table 13. Performance of Endoscopic Healing Index in Asymptomatic Patients With Crohn's Disease in Different Scenarios[a]

	No. of results per 1000 patients tested (95% CI)		
	Known endoscopic remission (prevalence 20%)	Unknown endoscopic remission status (prevalence 45%)	
Test result	EHI <20	EHI <20	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	188	423	TPs would be eligible to undergo treatment adjustment, which may decrease disease-related complications and morbidity, without being subject to risks and invasive testing with endoscopy.
FNs (patients incorrectly labeled as being in endoscopic remission or having only mild endoscopic activity, when actually they have moderate to severe endoscopic activity)	12	27	FNs would be falsely reassured, and may be at higher risk of disease complications/flare due to undertreatment.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	
TNs (patients correctly diagnosed as being in endoscopic remission or having only mild endoscopic activity)	424	292	TNs would be reassured and obviate the need for invasive testing with endoscopy, although they may need to undergo serial assessment of biomarker at periodic intervals.
FPs (patients incorrectly labeled as having moderate to severe endoscopic activity, when actually they are in endoscopic remission or have only mild endoscopic activity)	376	258	FPs may receive unnecessary testing (endoscopy) and/or treatment adjustment, and have avoidable anxiety, potential testing-, or treatment-related complications, and excessive resource utilization.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	

[a]Population/setting: adults with CD in symptomatic remission on stable maintenance therapy, with known endoscopic remission asymptomatic + known endoscopic remission in preceding 2–3 y, without change in clinical status and on stable therapy) with observed prevalence of active disease of 20%; with unknown endoscopic remission status (asymptomatic + unknown endoscopic remission status in preceding 2–3 y) with observed prevalence of endoscopically active disease of 45%. Sensitivity of EHI <20: 94%, 2 studies. Specificity of EHI <20: 53%, 2 studies. Reference test: colonoscopy.
[b]High unexplained heterogeneity, selective inclusion of studies reporting cutoffs.
[c]Very serious imprecision due to small number of studies, only involved in derivation-validation.

1. Biomarker-based treat-to-target strategy in CD: Treatment strategy trials, such as CALM, have demonstrated that incorporating biomarker assessment as part of the treat-to-target strategy is beneficial, especially in patients with active disease.[14] However, such treatment strategy trials have relied on a rigid set of prespecified criteria that would result in treatment escalation. There is a need for examination of various biomarker cutoff thresholds to guide therapy escalation and examination of the role of combination of biomarkers and role of biomarker-based treat-to-target strategy in asymptomatic patients, as well as potential harm of not dose escalating in the setting of mild biomarker abnormality to robustly inform biomarker-based treatment strategies. There have not been any studies comparing a biomarker-based strategy with an endoscopy-based strategy for assessment and monitoring of endoscopic remission. This was identified as a knowledge gap by the panel.

2. Magnitude of elevation of biomarkers: The guideline panel focused on examination of performance of biomarkers at commonly reported cutoffs that are widely used in clinical practice. Consequently, management recommendations could only be made based on whether the value was above the cutoff for that biomarker, but did not factor in the degree of abnormality. A single measurement demonstrating marked elevation of a biomarker may, for a given patient, carry a different prognostic implication than a more modest elevation. For example, in individuals with mild symptoms, fecal calprotectin >2500 μg/g may carry different implications for management than

Table 14. Performance of Endoscopic Healing Index In Patients With Symptomatically Active Crohn's Disease in Different Scenarios[a]

	No. of results per 1000 patients tested (95% CI)		
	Mild symptoms (prevalence 65%)	Moderate to severe symptoms (prevalence 80%)	
Test result	EHI >50	EHI >50	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	234	288	TPs would be eligible to undergo treatment adjustment, which may decrease disease-related complications and morbidity, without being subject to risks and invasive testing with endoscopy.
FNs (patients incorrectly labeled as being in endoscopic remission or having only mild endoscopic activity, when actually they have moderate to severe endoscopic activity)	416	512	FNs may be falsely reassured, undertreated, or mistreated (as not having CD flare), potentially leading to increased disease related complications and morbidity.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	
TNs (patients correctly diagnosed as being in endoscopic remission or having only mild endoscopic activity)	332	190	TNs would be reassured and obviate the need for invasive testing with endoscopy, although they may need to undergo serial assessment of biomarker at periodic intervals.
FPs (patients incorrectly labeled as having moderate to severe endoscopic activity, when actually they are in endoscopic remission or have only mild endoscopic activity)	18	10	FPs may undergo unnecessary treatment adjustment and have treatment-related complications.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	

[a]Population/setting: adults with CD with symptoms suggestive of active CD, mild symptoms (PRO2 8–13, or PRO3 13–21) with observed prevalence of endoscopically active disease of 65%; moderate to severe symptoms (PRO2 >13 or PRO3 >21) with observed prevalence of endoscopically active disease of 80%. Sensitivity of EHI >50, 36%, 2 studies. Specificity of EHI >50, 95%, 2 studies. Reference test: colonoscopy.
[b]High unexplained heterogeneity, selective inclusion of studies reporting cutoffs.
[c]Very serious imprecision due to small number of studies, only involved in derivation–validation.

fecal calprotectin of 251 μg/g.[36] There were insufficient data to guide nuanced decision making in this context. There are several novel biomarkers, including biomarker panels, of disease activity and prognosis that have been studied in research settings, but require more robust clinical validation before widespread adoption. The paucity of data on this was also identified as a knowledge gap by the panel, requiring further research.

3. Choice of treatment target: Consistent with existing clinical guidelines defining endoscopic remission in CD as SES-CD <3,[20] this guideline examined the diagnostic accuracy of biomarkers in determining either the presence or absence of inflammation at this threshold. Studies have demonstrated that a more rigorous treatment target of SES-CD score of 0 may be associated with better outcomes.[28] It is likely that biomarker performance will not be as robust against the more rigorous treatment target. There were limited data on performance of biomarkers against other treatment goals, such as histologic remission or transmural healing on radiologic assessment. This was identified as a knowledge gap. Conversely, some patients and physicians may elect to optimize therapy only for moderate to severe disease activity (SES-CD >6); the performance of the biomarker may be superior against that end point compared with that reported in this guideline.
4. Influence of disease location and extent: There was significant heterogeneity in extent and location of involved segment in CD; this may directly influence biomarker sensitivity and specificity, as well as its

Table 15. Performance of Endoscopic Healing Index in Asymptomatic Patients With Crohn's Disease in Surgically Induced Remission in Different Scenarios[a]

	No. of results per 1000 patients tested (95% CI)			
	EHI <20			
Test result	Low pretest probability (10%)	Intermediate pretest probability (30%)	High pretest probability (60%)	Comments
FNs (patients incorrectly labeled as being in endoscopic remission, when actually they have endoscopic activity)	18	54	108	FNs would be falsely reassured and may be at higher risk of disease complications/ progression and flare due to undertreatment and missed opportunity to treat.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	VERY LOW[b,c]	
FPs (patients incorrectly labeled as having endoscopic activity, when actually they are in endoscopic remission)	450	350	200	FPs may receive unnecessary testing (endoscopy) and/or treatment adjustment, and have avoidable anxiety, potential testing- or treatment-related complications, and excessive resource utilization.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	VERY LOW[b,c]	

[a]Population/setting: patients with CD in surgically induced remission, low pretest probability/likelihood of having endoscopic recurrence such as patients at low baseline risk of recurrence (no high-risk features) and on postoperative prophylaxis, 10%; intermediate pretest probability/likelihood of having endoscopic recurrence such as patients at low baseline risk of recurrence (no high-risk features), who are not on postoperative prophylaxis or patients at high baseline risk of recurrence who are receiving postoperative prophylaxis, 30%; high pretest probability/likelihood of having endoscopic recurrence, such as patients at high baseline risk of recurrence who are not receiving postoperative prophylaxis, 60%. Sensitivity of EHI >50, 82%, 1 study. Specificity of EHI >50, 50%, 1 study. Reference test: colonoscopy.
[b]High unexplained heterogeneity, selective inclusion of studies reporting cutoffs.
[c]Very serious imprecision due to small number of studies, only involved in derivation–validation.

accuracy. The correlation between symptoms and endoscopic activity may be weaker for small bowel CD, which would lead to a lower prevalence of endoscopically active disease for given symptoms. This would reduce the accuracy of the biomarker in individuals with symptomatic CD. There are no widely accepted validated scoring systems for endoscopic assessment of mucosal inflammation in CD involving the proximal small bowel in isolation (ie, beyond the reach of the colonoscope). Thus, the panel determined this to be a knowledge gap in the performance of biomarkers. Please see more detailed discussion of the impact of disease location on biomarker performance in the Key Considerations for Implementing These Recommendations in Clinical Practice section.

5. Biomarker performance in diverse populations: The panel recognized the lack of robust data in specific clinical situations, including mild CD, CD involving the J-pouch or in patients with an ostomy, and in geographically and ethnically diverse patient populations, where there exist few studies examining the role of biomarkers to date.

6. Comparison with other disease activity assessment modalities: Cross-sectional imaging is being used increasingly to define transmural healing in CD with growing use of computed tomography, magnetic resonance, and intestinal ultrasound–based assessments. Intestinal ultrasound, in particular, is attractive as a point-of-care test without radiation exposure and limited preparation. There were few studies comparing biomarker performance against these imaging-based assessments.

What Do Other Guidelines Say?

There has been limited discussion on the role of noninvasive biomarkers in the management of CD in clinical guidelines. The American College of Gastroenterology Society guidelines published in 2018 on the management of CD suggested fecal calprotectin and serum CRP may have an adjunctive role in assessing inflammation in patients with CD, but did not provide specific cutoffs or recommendations for use.[37] The European Crohn's and Colitis Organization and the European Society of Gastrointestinal and Abdominal Radiology guidelines on the diagnostic assessment of IBD recognized that in asymptomatic patients, elevated biomarkers of inflammation, mainly fecal calprotectin and CRP, may suggest imminent flare and recommended endoscopic or radiologic evaluation.[38] In patients with clinical response to medical therapy, the guidelines recommended evaluating for mucosal healing either via endoscopy or fecal

calprotectin. None of these guidelines discussed performance of specific cutoffs and downstream implications involved in decision making, which are critical to using these biomarkers in clinical practice. Similar to the current guideline, the AGA guideline for the use of biomarkers in UC suggested that a normal biomarker in asymptomatic patients or an elevated biomarker in those with moderate to severe symptoms can reliably rule out or rule in the presence of endoscopically active disease, respectively, thereby avoiding endoscopy solely for assessment of disease activity.[15] However, there are some key distinctions between the 2 guidelines. First, in CD, symptoms correlate less well with endoscopic activity. Thus, biomarker performance was only acceptable in asymptomatic individuals who had recently confirmed endoscopic remission; in those without recent endoscopic assessment, test performance was suboptimal, and the guideline suggests endoscopic assessment as the preferred strategy for assessing disease activity. Second, the weaker correlation between symptoms and endoscopic activity in CD also reduced the utility of biomarker measurement to infer disease activity in those with mild symptoms.

Plans for Updating This Guideline

Guidelines are living products. To remain useful, they need to be updated regularly as new information accumulates. This document will be updated when major new research is published. The need for update will be determined no later than 2026 and, if appropriate, we will update the guidelines to incorporate updated recommendations as new evidence, without duplicating or creating a new comprehensive guideline.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://doi.org/10.1053/j.gastro.2023.09.029.

References

1. Burisch J, Zhao M, Odes S, et al. The cost of inflammatory bowel disease in high-income settings: a *Lancet Gastroenterology & Hepatology* Commission. Lancet Gastroenterol Hepatol 2023;8:458–492.
2. Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet 2017;390:2769–2778.
3. Torres J, Mehandru S, Colombel JF, et al. Crohn's disease. Lancet 2017;389:1741–1755.
4. Lowe SC, Sauk JS, Limketkai BN, et al. Declining rates of surgery for inflammatory bowel disease in the era of biologic therapy. J Gastrointest Surg 2021;25:211–219.
5. Tsai L, Ma C, Dulai PS, et al. Contemporary risk of surgery in patients with ulcerative colitis and Crohn's disease: a meta-analysis of population-based cohorts. Clin Gastroenterol Hepatol 2021;19:2031–2045.e11.
6. Agrawal M, Colombel JF. Treat-to-target in inflammatory bowel diseases, what is the target and how do we treat? Gastrointest Endosc Clin North Am 2019;29:421–436.
7. Colombel J, Panaccione R, Bossuyt P, et al. A treat-to-target approach decreases the rate of CD-related adverse outcomes versus a clinical approach in patients with moderate to severely active Crohn's disease: data from CALM. United European Gastroenterol J 2017;5:A97–A98.
8. Colombel JF, Panaccione R, Bossuyt P, et al. Superior endoscopic and deep remission outcomes in adults with moderate to severe Crohn's disease managed with treat to target approach versus clinical symptoms: data from CALM. Gastroenterology 2017;152:S155.
9. Ungaro RC, Yzet C, Bossuyt P, et al. Deep remission at 1 year prevents progression of early Crohn's disease. Gastroenterology 2020;159:139–147.
10. Tse CS, Singh S, Valasek MA, et al. Prevalence and correlations of gastrointestinal symptoms with endoscopic and histologic mucosal healing in Crohn's disease. Am J Gastroenterol 2023;118:748–751.
11. Carman N, Tomalty D, Church PC, et al. Clinical disease activity and endoscopic severity correlate poorly in children newly diagnosed with Crohn's disease. Gastrointest Endosc 2019;89:364–372.
12. Lewis JD, Rutgeerts P, Feagan BG, et al. Correlation of stool frequency and abdominal pain measures with Simple Endoscopic Score for Crohn's Disease. Inflamm Bowel Dis 2020;26:304–313.
13. Froslie KF, Jahnsen J, Moum BA, et al. Mucosal healing in inflammatory bowel disease: results from a Norwegian population-based cohort. Gastroenterology 2007;133:412–422.
14. Colombel J-F, Panaccione R, Bossuyt P, et al. Effect of Tight Control Management on Crohn's Disease (CALM): a multicentre, randomised, controlled phase 3 trial. Lancet 2017;390:2779–2789.
15. Singh S, Ananthakrishnan AN, Nguyen NH, et al. AGA Clinical Practice Guideline on the role of biomarkers for the management of ulcerative colitis. Gastroenterology 2023;164:344–372.
16. Schunemann HJ, Oxman AD, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ 2008;336:1106–1110.
17. Sands BE. Biomarkers of inflammation in inflammatory bowel disease. Gastroenterology 2015;149:1275–1285.e2.
18. Reitsma JB, Glas AS, Rutjes AW, et al. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005;58:982–990.
19. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–560.
20. Vuitton L, Marteau P, Sandborn WJ, et al. IOIBD technical review on endoscopic indices for Crohn's disease clinical trials. Gut 2016;65:1447–1455.
21. Peyrin-Biroulet L, Sandborn W, Sands BE, et al. Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE): determining therapeutic goals for treat-to-target. Am J Gastroenterol 2015;110:1324–1338.

22. 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 (IOIBD): Determining Therapeutic Goals for Treat-to-Target strategies in IBD. Gastroenterology 2021;160:1570–1583.
23. Barsky M, Meserve J, Le H, et al. Understanding determinants of patient preferences between stool tests and colonoscopy for the assessment of disease activity in inflammatory bowel disease. Dig Dis Sci 2021; 66:2564–2569.
24. Dulai PS, Singh S, Jairath V, et al. Prevalence of endoscopic improvement and remission according to patient-reported outcomes in ulcerative colitis. Aliment Pharmacol Ther 2020;51:435–445.
25. Nguyen GC, Loftus EV Jr, Hirano I, et al. American Gastroenterological Association Institute Guideline on the management of Crohn's disease after surgical resection. Gastroenterology 2017;152:271–275.
26. Rutgeerts P, Geboes K, Vantrappen G, et al. Predictability of the postoperative course of Crohn's disease. Gastroenterology 1990;99:956–963.
27. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–536.
28. Yzet C, Diouf M, Le Mouel J-P, et al. Complete endoscopic healing associated with better outcomes than partial endoscopic healing in patients with Crohn's disease. Clin Gastroenterol Hepatol 2020;18:2256–2261.
29. Simon EG, Wardle R, Thi AA, et al. Does fecal calprotectin equally and accurately measure disease activity in small bowel and large bowel Crohn's disease?: a systematic review. Intest Res 2019;17:160–170.
30. Danese S, Vermeire S, D'Haens G, et al. Treat to target versus standard of care for patients with Crohn's disease treated with ustekinumab (STARDUST): an open-label, multicentre, randomised phase 3b trial. Lancet Gastroenterol Hepatol 2022;7:294–306.
31. Barnes EL, Lightner AL, Regueiro M. Perioperative and postoperative management of patients with Crohn's disease and ulcerative colitis. Clin Gastroenterol Hepatol 2020;18:1356–1366.
32. Lee KE, Cantrell S, Shen B, et al. Post-operative prevention and monitoring of Crohn's disease recurrence. Gastroenterol Rep (Oxf) 2022;10:goac070.
33. Regueiro M, Velayos F, Greer JB, et al. American Gastroenterological Association Institute technical review on the management of Crohn's disease after surgical resection. Gastroenterology 2017;152:277–295.e3.
34. D'Haens G, Kelly O, Battat R, et al. Development and validation of a test to monitor endoscopic activity in patients with Crohn's disease based on serum levels of proteins. Gastroenterology 2020;158:515–526.e10.
35. Hamilton AL, De Cruz P, Wright EK, et al. Non-invasive serological monitoring for Crohn's disease postoperative recurrence. J Crohns Colitis 2022;16:1797–1807.
36. Hyams JS, Davis Thomas S, Gotman N, et al. Clinical and biological predictors of response to standardised paediatric colitis therapy (PROTECT): a multicentre inception cohort study. Lancet 2019;393:1708–1720.
37. Lichtenstein GR, Loftus EV, Isaacs KL, et al. ACG Clinical Guideline: management of Crohn's disease in adults. Am J Gastroenterol 2018;113:481–517.
38. Maaser C, Sturm A, Vavricka SR, et al. ECCO-ESGAR guideline for diagnostic assessment in IBD part 1: initial diagnosis, monitoring of known IBD, detection of complications. J Crohns Colitis 2019;13:144–164.

Correspondence
Address correspondence to: Chair, Clinical Guidelines Committee, American Gastroenterological Association, National Office, 4930 Del Ray Avenue, Bethesda, Maryland 20814. e-mail: clinicalpractice@gastro.org.

Acknowledgments
The authors would like to thank Ms Caitlin Bakker, University of Minnesota Libraries, for designing and conducting the literature search.

Conflicts of interest
These authors disclose the following: Jeremy Adler receives research grants from Janssen Research & Development. Benjamin L. Cohen receives the following financial support: consulting fees from AbbVie, Celgene-Bristol Myers Squibb, Lilly, Pfizer, Sublimity Therapeutics, Takeda, TARGET RWE; CME Companies: Cornerstones, Vindico; Speaking: AbbVie; Educational Grant: Pfizer. Siddharth Singh's institution has received research grants from Pfizer and AbbVie, and he has received personal fees from Pfizer (for ad hoc grant review). The remaining authors disclose no conflicts. A full list of conflicts active at the time of guideline development can be accessed at AGA's National Office in Bethesda, MD.

Funding
These guidelines were fully funded by the AGA Institute. Dr Ananthakrishnan is supported by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grants R21DK127227 and R01 DK127171, in addition to grants from the Leona M and Harry B Helmsley Charitable Trust and the Chleck Family Foundation. Dr Siddique is supported by NIDDK grant K08DK120902. Dr Singh is supported by NIDDK grants K23DK117058 and R03DK129631.