Gastroenterology 2023;164:344–372

GUIDELINES

AGA Clinical Practice Guideline on the Role of Biomarkers for the Management of Ulcerative Colitis

Siddharth Singh,[1,2,*] **Ashwin N. Ananthakrishnan**,[3,4,*] Nghia H. Nguyen,[1] Benjamin L. Cohen,[5] Fernando S. Velayos,[6] Jennifer M. Weiss,[7] Shahnaz Sultan,[8,9] Shazia M. Siddique,[10,11] Jeremy Adler,[12,13] and Karen A. Chachu,[14] on behalf of the AGA Clinical Guidelines Committee

[1]Division of Gastroenterology and Hepatology, Department of Medicine, University of California San Diego, La Jolla, California; [2]Division of Biomedical Informatics, Department of Medicine, University of California San Diego, La Jolla, California; [3]Division of Gastroenterology, Massachusetts General Hospital, Boston, Massachusetts; [4]Harvard Medical School, Boston, Massachusetts; [5]Division of Gastroenterology, Hepatology, and Nutrition, Digestive Disease and Surgery Institute, Cleveland Clinic, Cleveland, Ohio; [6]Division of Gastroenterology, Kaiser Permanente Medical Group, San Francisco, California; [7]Division of Gastroenterology and Hepatology, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin; [8]Division of Gastroenterology, Hepatology, and Nutrition, University of Minnesota, Minneapolis, Minnesota; [9]Veterans Affairs Healthcare System, Minneapolis, Minnesota; [10]Division of Gastroenterology, Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania; [11]Center for Evidence-Based Practice, University of Pennsylvania Health System, Philadelphia, Pennsylvania; [12]Division of Pediatric Gastroenterology, CS Mott Children's Hospital, Michigan Medicine, University of Michigan, Ann Arbor, Michigan; [13]Susan B. Meister Child Health Evaluation and Research Center, University of Michigan, Ann Arbor, Michigan; and [14]Division of Gastroenterology, Department of Medicine, Duke University, Durham, North Carolina

BACKGROUND & AIMS: Biomarkers are used frequently for noninvasive monitoring and treatment decision making in the management of patients with ulcerative colitis (UC). This American Gastroenterological Association (AGA) guideline is intended to support practitioners in decisions about the use of biomarkers for the management of UC. **METHODS:** A multidisciplinary panel of content experts and guideline methodologists used the Grading of Recommendations Assessment, Development and Evaluation framework to prioritize clinical questions, identify patient-centered outcomes, and conduct an evidence synthesis on the clinical performance of serum C-reactive protein (CRP), fecal calprotectin, and fecal lactoferrin as biomarkers of disease activity in patients with established UC in symptomatic remission or with active symptoms. The guideline panel used the Evidence-to-Decision framework to develop recommendations for the use of biomarkers for monitoring and management of UC and provided implementation considerations for clinical practice. **RESULTS:** The guideline panel made 7 conditional recommendations. In patients with UC in symptomatic remission, the panel suggests the use of a biomarker- and symptom-based monitoring strategy over a symptom-based monitoring strategy. For patients in symptomatic remission, the panel suggests using fecal calprotectin <150 μg/g, normal fecal lactoferrin, and/or normal CRP to rule out active inflammation and avoid routine endoscopic assessment of disease. In patients with UC with moderate to severe symptoms, the panel suggests using fecal calprotectin >150 μg/g, elevated fecal lactoferrin, or elevated CRP to inform treatment decisions and avoid routine endoscopic assessment of disease. However, in patients in symptomatic remission but elevated biomarkers, and in patients with moderate to severe symptoms with normal biomarkers, the panel suggests endoscopic assessment of disease to inform treatment decisions. In patients with UC with mild symptoms, the panel suggests endoscopic assessment of disease activity to inform treatment decisions. The panel identified the use of a biomarker-based monitoring strategy over an endoscopy-based monitoring strategy as a knowledge gap. The panel also proposed key implementation considerations for optimal use of biomarkers, and identified areas for future research. **CONCLUSIONS:** In patients with UC, noninvasive biomarkers, including fecal calprotectin, fecal lactoferrin, and serum CRP can inform disease monitoring and management.

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

Inflammatory bowel diseases comprising Crohn's disease (CD) and ulcerative colitis (UC) are estimated to affect more than 7 million individuals worldwide.[1,2] UC is characterized by periods of relapsing–remitting activity, resulting in considerable morbidity.[3] Up to 1 in 5 patients with UC may undergo definitive surgery in the form of a total colectomy, often for medically refractory disease.[4] The direct and indirect costs attributable to UC are considerable and continue to increase.[5]

There has been a paradigm shift in the management of UC over the past 2 decades. The therapeutic target has

***Authors share co-first authorship.**

Abbreviations used in this paper: **AGA, American Gastroenterological Association; CD, Crohn's disease; CRP, C-reactive protein; FN, false negative; FP, false positive; GRADE, Grading of Recommendations Assessment, Development and Evaluation; MES, Mayo Endoscopic Score; PICO, patients, intervention, comparator, and outcome; RBS, rectal bleeding score; RCT, randomized controlled trial; SFS, stool frequency score; TN, true negative; TP, true positive; UC, ulcerative colitis.**

Most current article

0016-5085/$36.00
https://doi.org/10.1053/j.gastro.2022.12.007

shifted from symptom resolution alone to a combination of symptomatic and endoscopic remission (Mayo Endoscopic Score [MES] 0 or equivalent) or improvement (MES 0 or 1).[6,7] Achievement of endoscopic remission or improvement is associated with superior outcomes, including lower risk of relapse, need for corticosteroids, hospitalizations, colectomy, and colorectal neoplasia.[8,9] Although there are no randomized controlled trials (RCTs) in UC, indirect evidence from treatment strategy intervention trials in CD, such as the CALM study, in which a tight control strategy based on a combination of symptoms and biomarkers was more effective than a usual care strategy targeting symptoms alone in achieving deep remission, which, in turn, was associated with lower risk of disease progression and complications, surgery, and hospitalization.[10,11]

Most RCTs have relied on endoscopic evaluation to confirm resolution of bowel inflammation. Similarly, in clinical practice, endoscopic assessment of bowel inflammation after initiation of therapy is performed in 45%–70% of patients.[12,13] Despite the fact that early proactive assessment of bowel inflammation is associated with superior long-term outcomes, there is significant variability in utilization.[12] Moreover, in routine clinical practice, repeated endoscopic assessment is invasive, expensive, and may be impractical. There is an important need for understanding how noninvasive biomarkers may serve as accurate and reliable surrogates for endoscopic assessment of inflammation and whether they can be more readily implemented in a UC care pathway. Finally, patients with UC may prefer alternative noninvasive tests, such as biomarkers, over endoscopy, although this preference varies depending on the diagnostic and prognostic performance of biomarkers.[14]

Objective

The objective of this guideline was to provide guidance about the use of well-established and commonly available biomarkers as surrogate tests for endoscopic assessment of disease or in longitudinal monitoring of patients with an established diagnosis of UC primarily in the ambulatory setting. Predictive biomarkers for assessment of hospitalized patients with acute severe colitis is beyond the scope of this guideline. This guideline focuses on the following biomarkers: serum C-reactive protein (CRP), fecal calprotectin, and fecal lactoferrin. Laboratory evaluation of diarrhea in patients with suspected UC is discussed elsewhere.[15,16] The role of biomarkers in patients with CD will be discussed in a subsequent guideline.

Target Audience

The target audience for this guideline includes primary care and gastroenterology health care professionals, patients, and policy makers. This guideline is not intended to impose a standard of care, but rather provide the basis for rational informed decisions for patients and health care professionals. Statements regarding the underlying values and preferences, as well as qualifying comments accompanying each recommendation, should never be omitted when quoting or translating recommendations from this guideline. Recommendations provide guidance for typical patients with UC; no recommendation can consider all unique circumstances that must be accounted for when making recommendations for individual patients. However, discussions about benefits and harms can be used for shared decision making, especially for conditional recommendations in which specific tradeoffs and patient values 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).[17] The development of this guideline was fully funded by the AGA Institute.

Guideline Panel Composition and Conflict of Interest

Members of the guideline and evidence synthesis panel were selected on the basis of their clinical and methodological expertise, after undergoing a vetting process that required disclosing all conflicts of interest. The evidence synthesis panel consisted of 2 content experts with expertise in UC (Ashwin N. Ananthakrishnan, Jeremy Adler), a senior guideline methodologist with expertise in evidence synthesis and GRADE (Siddharth Singh), and 2 junior guideline methodologists (Nghia H. Nguyen, Shazia M. Siddique). The guideline panel consisted of a multidisciplinary panel that included a general gastroenterologist (Jennifer M. Weiss), gastroenterologists with expertise in inflammatory bowel diseases (Karen A. Chachu, Benjamin L. Cohen, Fernando S. Velayos), and guideline methodologists (Siddharth Singh, Nghia H. Nguyen, Shazia M. Siddique, Siddharth Sultan). A patient representative also participated in the development of the guideline recommendations. 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 (Karen A. Chachu) and co-chair and senior methodologist (Siddharth Singh) had no conflicts of interest. No guideline panel member was excused from participation in the process owing to disqualifying conflict. A full list of conflicts can be accessed at AGA's National Office in Bethesda, MD.

Scope

Biomarkers are defined biological molecules that are quantifiable in tissue or body fluid (blood, stool, and urine) that represent an underlying biological disease process. Various biomarkers have been investigated extensively in UC for several outcomes, including prediction of onset; establishing diagnosis; assessing disease activity; prognosticating natural history, including likelihood of colectomy; and assessing postcolectomy outcomes.[18] However, most of these studies have been small, lack replication, and use markers that are not readily available outside of a research setting. For inclusion, we

required the biomarker to be both readily measurable in a tissue or body fluid compartment, widely available for commercial use, and used routinely in day-to-day clinical practice for assessing disease activity or providing actionable prognostic information longitudinally. Based on these criteria, we focused on serum CRP, fecal calprotectin, and fecal lactoferrin.

Formulation of Clinical Questions

Through an iterative process, the guideline and evidence synthesis panels developed focused clinical questions deemed relevant for clinical practice that the guideline would address, related to the diagnostic performance and utility of commonly used serum and stool biomarkers in patients with established UC. From these focused questions, well-defined statements in terms of patients, intervention, comparator, and outcome (PICO) were defined, and these formed the framework for formulating the study inclusion and exclusion criteria and guided the literature search. The AGA Governing Board approved the final set of questions and statements in October 2021. The final focused questions and PICO questions are shown in Table 1.

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). To ensure that recent studies were not missed, searches were updated before external review. The search was limited to English language and human subjects. The bibliography 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 ongoing studies.

Study Selection, Data Abstraction, and Statistical Analysis

We included RCTs or observational studies of diagnostic accuracy that met the following inclusion criteria: performed in patients with UC, provided adequate description of biomarker (ie, CRP, fecal calprotectin, and/or fecal lactoferrin) with cutoff corresponding to detection of moderate to severe endoscopic activity (corresponding to MES 2 or 3), with lower endoscopy as reference standard, and provided sufficient data to allow estimation of diagnostic accuracy of biomarker (ie, sensitivity and specificity) for detection of endoscopic activity. For these questions, we preferentially chose cutoffs most commonly used in clinical practice rather than focusing on optimized cutoffs identified in individual studies. The cutoffs were as follows: C-reactive protein: 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 fecal lactoferrin: normal or elevated based on laboratory cutoff.

From each study, we abstracted data on study, patient, biomarker, outcome, and test performance data (ie, sensitivity, specificity, prevalence of outcome of interest in study, to impute numbers of true-positive [TP], true-negative [TN], false-positive [FP], and false-negative [FN] results). The paired values of sensitivity and specificity were pooled using a bivariate regression random-effects model proposed by Reitsma et al[19] 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.[20]

Outcomes of Interest and Illustrative Clinical Scenarios

For PICO questions focusing on biomarker cutoffs to either detect or rule out moderate to severe endoscopic activity, the preferred outcome was direct consequences on patient-important outcomes (ie, implications of TP, FP, TN, and FN results for patients). However, none of the studies assessed these outcomes directly, and hence, we used TP, FP, TN, and FN rates as surrogate outcomes and inferred downstream consequences on patient-important outcomes.

For questions focusing on ruling out moderate to severe endoscopic activity, our outcome was minimizing rates of FN (ie, patients incorrectly being labeled as being in remission, when they actually have moderate to severe endoscopic inflammation) to a level $<5\%$ in general, preferably lower, with reasonable rates of TP, FP, and TN (Supplementary Figure 1). For questions focusing on detecting moderate to severe endoscopic activity, our outcome was minimizing rates of FP (ie, patients incorrectly labeled as having moderate to severe endoscopic inflammation when their disease is actually in remission) (Supplementary Figure 2). These thresholds of 5% FN and FP rate were also consistent with patient preference for choosing stool-based biomarkers over endoscopic assessment for monitoring inflammation.[14]

Although sensitivity and specificity are agnostic of disease prevalence, overall TP, FP, TN, and FN rates are highly dependent on pretest probability. We derived illustrative prevalence of moderate to severe endoscopic activity based on a combination of rectal bleeding score (RBS) and stool frequency score (SFS), 2 of the most commonly used patient-reported outcomes, derived from the MES (Supplementary Table 2). Prevalence of moderate to severe endoscopic activity (MES 2 or 3) and endoscopic improvement (MES 0 or 1), for different combinations of RBS and SFS, at varying time points after treatment initiation/adjustment were derived from existing literature based on individual participant data from phase 2 and 3 clinical trial programs of biologic agents and small molecule inhibitors in patients with moderate to severely active UC.[21]

For our analysis, we used the following 3 illustrative scenarios:

1. Low pretest probability of having moderate to severe inflammation: These include patients with asymptomatic UC (no rectal bleeding and normal to mild increase in stool frequency, RBS 0, and SFS 0 or 1), on stable maintenance therapy, or having recently achieved symptomatic remission after treatment adjustment. The expected prevalence of moderate to severe endoscopic inflammation in these patients is approximately 15%.

2. Intermediate pretest probability of having moderate to severe inflammation: These include patients with mild symptoms of UC, such as infrequent rectal bleeding (RBS 0 or 1) and/or increased stool frequency (SFS 2 or 3). The prevalence of moderate to severe endoscopic inflammation in these patients is approximately 50%.

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

Question no.	Focused question	Patients	Intervention (threshold?)	Comparator	Outcome
Patients with UC in symptomatic remission					
1	In patients with UC in symptomatic remission, is interval biomarker-based monitoring superior to symptom-based monitoring to improve long-term outcomes?	Patients with established UC in symptomatic remission	Interval biomarker-based monitoring	Interval symptom-based monitoring	Maintaining clinical remission at 12 mo and beyond
2	In patients with UC in symptomatic remission, at what (A) fecal calprotectin, (B) fecal lactoferrin, and (C) serum C-reactive protein cutoff can we accurately rule out active inflammation, obviating routine endoscopic assessment?	Patients with established UC in symptomatic remission, or with mild symptoms in whom fecal calprotectin, fecal lactoferrin and serum CRP was measured	Fecal calprotectin <50 μg/g, <150 μg/g, or <250 μg/g Normal fecal lactoferrin (<7.25 μg/g) Normal CRP (<5 mg/L)	Fecal calprotectin >50 μg/g, >150 μg/g, or >250 μg/g Elevated fecal lactoferrin (>7.25 μg/g) Elevated CRP (>5 mg/L)	For detection of endoscopic inflammation Benefits: 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 UC					
3	In patients with symptomatically active UC, is an evaluation strategy that combines biomarkers and symptoms superior to symptom-based evaluation for making treatment adjustments?	Patients with symptomatically active UC	Biomarker-based evaluation	Symptom-based evaluation	For detection of endoscopic inflammation Benefits: 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

Question no.	Focused question	Patients	Intervention (threshold?)	Comparator	Outcome
4	In patients with symptomatically active UC, at what (A) fecal calprotectin, (B) fecal lactoferrin, and (C) serum CRP cutoff can we accurately diagnose active inflammation, obviating routine endoscopic assessment?	Patients with established UC with typical symptoms suggestive of flare or mild symptoms in whom fecal calprotectin, fecal lactoferrin, and serum CRP were measured	Fecal calprotectin >50 μg/g, >150 μg/g, or >250 μg/g Elevated fecal lactoferrin (>7.25 μg/g) Elevated CRP (>5 mg/L)	Fecal calprotectin <50 μg/g, <150 μg/g or <250 μg/g Normal fecal lactoferrin (<7.25 μg/g) Normal CRP (<5 mg/L)	For detection of endoscopic inflammation Benefits: 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)
Treat-to-target strategies for UC					
5	In patients with established UC, is interval biomarker-based monitoring strategy superior to interval endoscopy-based monitoring strategy to improve long-term outcomes?	Patients with UC in symptomatic remission	Interval biomarker-based monitoring	Interval endoscopy-based monitoring	Maintaining clinical remission at 12 mo and beyond

3. High pretest probability of having moderate to severe inflammation: These include patients with typical symptoms of active UC with frequent rectal bleeding and significant increase in stool frequency (RBS 2 or 3 and SFS 2 or 3). The prevalence of moderate to severe endoscopic inflammation in these patients is approximately 85%.

Consequences of Diagnostic Test Results on Patient-Important Outcomes

Corresponding to each possible outcome (TP, FP, TN, and FN), presumed downstream consequences on patient-important outcomes were considered. In using specific biomarkers either as a test replacement or triage strategy, health care providers and patients need to be aware of test performance and be comfortable with potential FN and FP rates with related downstream consequences. Such downstream consequences of test results for each PICO statement and scenario are discussed in detail in each evidence profile.

A premeeting questionnaire was administered to all members of the guideline panel and evidence synthesis panel to determine their *a priori* maximal tolerable FN rate and FP rate for each PICO (ie, what level of FN and FP rate would they be willing to accept for a particular test for their patient). As the maximally tolerable rates of FN and FP for any diagnostic strategy is highly context-sensitive, we devised different clinical scenarios with corresponding downstream consequences for each PICO to arrive at fully contextualized estimates of FN and FP thresholds.

Certainty of the Evidence

We rated the certainty of evidence using the GRADE approach for diagnostic tests and strategies.[17] In this approach, all evidence from RCTs (comparing different diagnostic tests or cutoffs of 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).[22]
- Indirectness (deemed present if there are 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 are 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 improvement for moderately to severe endoscopic activity were not prespecified, but primarily obtained 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).
- 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 21, 2022 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 recommendation are provided for each clinical question. As per GRADE methodology, recommendations are labeled as "strong" or "conditional." The phrase "we recommend" indicates strong recommendations and "we suggest" indicates conditional recommendations and provide the suggested interpretation of strong and weak recommendations for patients, clinicians, and health care policy makers.

Review Process

This guideline was submitted for public comment and internal review and was approved by the AGA Governing Board.

Discussion of Recommendations

A summary of all the recommendations is provided in Table 2 and each recommendation is discussed below. Key implementation considerations when considering using biomarkers in UC are discussed below and are summarized in Table 3.

Key Considerations for Implementing These Recommendations in Clinical Practice

In using a biomarker as a test replacement or triage strategy, it is critical to have a framework for understanding how each possible test result (TP, FP, TN, and FN) is associated with downstream consequences and impact on patient-important outcomes (Table 4).

The recommendations presented in this guideline are intended to provide a framework for incorporating biomarkers in the management pathway of patients with UC to inform treatment decisions. It is also critical to understand the limitations in interpreting these tests in various settings.

1. Considerations of test performance and specificity of biomarkers: The serum and fecal biomarkers in this guideline are not specific for UC activity. Serologic biomarkers, such as CRP, may be influenced by concurrent systemic illnesses, as well as other co-existing inflammatory diseases. Fecal markers, although more specific for intestinal inflammation, are not specific for UC disease activity and may be elevated in other inflammatory diseases of the gut, including infectious gastroenteritis and drug-induced colitis.[23] Thus, one should also consider simultaneous evaluation for enteric pathogens in patients with UC who present with gastrointestinal symptoms. Gastrointestinal infections are detected in approximately one-third of patients with UC presenting with gastrointestinal symptoms.[24,25]

Table 2. Executive Summary of Recommendations

Recommendation	Strength of recommendation	Certainty of evidence
Patients with ulcerative colitis in symptomatic remission		
In patients with UC in symptomatic remission, the AGA suggests a monitoring strategy that combines biomarkers and symptoms, rather than symptoms alone. Comment: Patients who place high value on avoiding burden of biomarker testing, over a potentially higher risk of flare or overtreatment, may reasonably choose interval symptom-based monitoring. Implementation considerations: • Interval biomarker monitoring may be performed every 6–12 mo. • Fecal biomarkers (fecal calprotectin or fecal lactoferrin) may be optimal for monitoring and may be particularly useful in patients where biomarkers have historically correlated with endoscopic disease activity. • A biomarker-based monitoring strategy, especially using stool-based tests, however, may be inconvenient and elevated biomarkers in otherwise asymptomatic individuals may lead to high patient anxiety. • It is important to think about the downstream consequences of testing and associated costs. The optimal management strategy in cases of discrepancy between symptoms and biomarkers is unclear and would generally trigger additional endoscopic testing for confirmation or repeat biomarker testing.	Conditional	Moderate
In patients with UC in symptomatic remission, the AGA suggests using fecal calprotectin <150 μg/g, normal fecal lactoferrin, or normal CRP to rule out active inflammation and avoid routine endoscopic assessment of disease activity Implementation considerations: • In patients who have recently achieved symptomatic remission after treatment adjustment in the preceding 1–3 mo, fecal calprotectin $<$50 μg/g may be preferred over $<$150 μg/g to detect endoscopic improvement (MES 0 or 1). • Normal CRP may be less informative to rule out moderate to severe active endoscopic inflammation in patients with UC in symptomatic remission, particularly in patients who have recently achieved symptomatic remission after treatment adjustment. However, if CRP was elevated at time of initial flare, then normalization of CRP may suggest endoscopic improvement (MES 0 or 1).	Conditional	Low (fecal calprotectin and fecal lactoferrin) to very low (CRP)
In patients with UC in symptomatic remission but elevated stool or serum markers of inflammation (fecal calprotectin >150 μg/g, elevated fecal lactoferrin, elevated CRP), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. Implementation consideration: In patients with UC in symptomatic remission but elevated biomarkers of inflammation, repeat measurement of biomarkers (in 3–6 mo) may be a reasonable alternative to endoscopic assessment. However, if biomarkers are elevated on repeat evaluation, then endoscopic assessment may be warranted.	Conditional	Very low

Table 2. Continued

Recommendation	Strength of recommendation	Certainty of evidence
Patients with symptomatically active ulcerative colitis		
In patients with symptomatically active UC, the AGA suggests an evaluation strategy that combines biomarkers and symptoms, rather than symptoms alone, to inform treatment adjustments. Comment: Patients, particularly those with severe symptoms, who place high value in avoiding burden of biomarker testing, over a potentially higher risk of inappropriate overtreatment, may reasonably choose symptom-informed treatment decisions.	Conditional	Low
In patients with UC with moderate to severe symptoms suggestive of flare, the AGA suggests using fecal calprotectin >150 μg/g, elevated fecal lactoferrin, or elevated CRP to rule in active inflammation and inform treatment adjustment and avoid routine endoscopic assessment solely for establishing presence of active disease. Comment: Patients who place greater value in confirming inflammation, particularly when making significant treatment decisions (such as starting or switching immunosuppressive therapies) and lesser value on the inconvenience of endoscopy, may choose to pursue endoscopic evaluation before treatment adjustment.	Conditional	Moderate (CRP), low (fecal calprotectin) to very low (fecal lactoferrin)
In patients with UC with mild symptoms, with elevated stool or serum markers of inflammation (fecal calprotectin >150 μg/g, elevated fecal lactoferrin, or elevated CRP), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. Implementation consideration: In patients with UC who underwent recent adjustment of treatment in response to moderate to severe symptomatic flare, and now have mild residual symptoms, elevated stool or serum markers of inflammation may be used to inform treatment adjustments (such as dose adjustments of therapy).	Conditional	Very low
In patients with UC with mild symptoms, with normal stool or serum markers of inflammation (fecal calprotectin <150 μg/g, normal fecal lactoferrin, normal CRP), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. Implementation consideration: • In patients with UC with mild symptoms (eg, slight increase in stool frequency and/or infrequent rectal bleeding), it may be reasonable to proceed directly with endoscopic assessment rather than testing biomarkers of inflammation. • In patients with UC with mild symptoms and normal biomarkers of inflammation who prefer to avoid endoscopic assessment or empiric treatment escalation, repeat measurement of biomarkers (in 3–6 mo) may be a reasonable alternative.	Conditional	Very low
Treat-to-target strategies for ulcerative colitis		
In patients with UC, 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

Table 3. Summary of Key Considerations When Using Biomarkers for Monitoring in Ulcerative Colitis

	Considerations
1	**Considerations of test performance and specificity of biomarkers:** CRP, fecal calprotectin, and fecal lactoferrin may be elevated because of nonintestinal sources of infection or inflammation. In patients with UC 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:** Biomarkers of inflammation have no role in dysplasia detection and surveillance and ruling out cytomegalovirus colitis, and endoscopic evaluation is the main strategy for evaluating these. Endoscopic evaluation may be useful for prognostication in patients hospitalized with acute severe UC.
3	**Association between treatment target and biomarker performance:** Test performance of all biomarkers in this guideline reflect their ability to rule out moderate to severe endoscopic inflammation (MES 2 or 3 [or equivalent]). Biomarkers may be suboptimal for detecting more rigorous treatment targets such as endoscopic remission (MES 0) or histologic remission. Biomarkers may also be suboptimal in detecting the presence of mild endoscopic activity (MES 1) in patients with mild symptoms.
4	**Influence of disease extent on biomarker performance:** Biomarkers may be less accurate in detecting endoscopic inflammation in patients with ulcerative proctitis or limited segmental disease.
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 variations 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 active disease and remission).

CRP, fecal calprotectin, and fecal lactoferrin may be elevated because of nonintestinal sources of infection or inflammation. In patients with UC 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: This guideline includes recommendations on the use of biomarkers as a replacement strategy for endoscopy in individuals with moderate to severe UC; however, endoscopy is frequently performed for other indications, for example, detection and surveillance of dysplasia in patients with long-standing UC.[26] Similarly, in patients with severe UC, particularly those who are refractory to corticosteroids, endoscopic assessment may be warranted to rule out cytomegalovirus colitis. In patients hospitalized with acute severe UC, endoscopic evaluation may be helpful to prognosticate and inform treatment.[27]

Biomarkers of inflammation have no role in dysplasia detection and surveillance, and ruling out cytomegalovirus colitis, and endoscopic evaluation is the main strategy for evaluating these.

Endoscopic evaluation may be useful for prognostication in patients hospitalized with acute severe UC.

3. Association between treatment target and biomarker performance: Current treatment guidelines recommend a target of endoscopic improvement (MES 0 or 1), although more updated consensus statements recommend a target of endoscopic remission (MES 0, or equivalent).[6,7] In a meta-analysis of 15 eligible studies, an MES of 0 was associated with a lower risk of clinical relapse (odds ratio, 0.33; 95% CI, 0.26–0.43) compared with an MES of 1.[28] Furthermore, histologic healing may also be a superior therapeutic goal; persistent histologic activity, even in the setting of endoscopically healed mucosa, is associated with a higher risk of relapse.[28,29] In this guideline, we focused on the accuracy of biomarkers to detect moderate to severe endoscopic inflammation (MES 2 or 3) because, currently in clinical practice, these are widely accepted triggers for treatment adjustment to achieve a conventional treatment target of MES 0 or 1. Diagnostic performance of a combination of symptoms and biomarkers to detect more rigorous end points, such as MES 0 or histologic remission, was not assessed in this guideline, but are likely to have

Table 4. Consequences of Diagnostic Test Results on Patient-Important Outcomes

Variable	Outcome
TP	Patients correctly diagnosed as having moderate to severe endoscopic activity 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.
FP	Patients incorrectly labeled as having moderate to severe endoscopic activity, 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.
TN	Patients correctly diagnosed as being in endoscopic remission or having only mild endoscopic activity 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.
FN	Patients incorrectly labeled as being in endoscopic remission or having only mild endoscopic activity, 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.

inferior performance, given differences in pretest probability. On a related note, a subset of patients with symptomatically active UC, generally with mild symptoms, may have mild inflammation on endoscopy (MES 1). The performance of biomarkers to specifically distinguish endoscopic remission (MES 0) vs mild endoscopic activity (MES 1) is limited.

Test performance of all biomarkers in this guideline reflect their ability to rule out moderate to severe endoscopic inflammation (MES 2 or 3 [or equivalent]). Biomarkers may be suboptimal for detecting more rigorous treatment targets, such as endoscopic remission (MES 0) or histologic remission. Biomarkers may also be suboptimal in detecting the presence of mild endoscopic activity (MES 1) in patients with mild symptoms.

4. Influence of disease extent on performance of fecal biomarkers: Elevation of fecal biomarkers such as calprotectin and lactoferrin may be influenced by the extent and location of inflamed surface. In a prospective study of patients with UC undergoing ileocolonoscopy, fecal calprotectin values demonstrated a stronger correlation with the extent of inflamed surface ($r = 0.86$) than region of maximal severity, independent of the severity of inflammation ($r = 0.79$).[30] In studies examining specific disease locations, the performance of fecal calprotectin in identifying active disease (MES 1–3) was weaker for proctitis ($r = 0.54$) compared with either left-sided colitis ($r = 0.75$) or extensive colitis ($r = 0.78$).[31] In other studies, fecal calprotectin was unable to accurately identify active disease in the setting of isolated proctitis in comparison with disease more extensively involving the colon; although in some studies, fecal calprotectin has demonstrated value in serially monitoring response to suppository treatment in isolated UC proctitis.[32,33] In the studies included in this literature review, the median proportion of patients with proctitis was 17% (interquartile range, 6.5%–27%) among the studies that reported disease extent.

Fecal biomarkers may be less accurate in detecting endoscopic inflammation in patients with ulcerative proctitis or limited segmental disease.

5. Interpreting biomarker performance for low-risk vs high-risk treatment adjustments: The acceptable threshold for performance of the biomarkers may differ based on the absolute and/or perceived cost and risk of the proposed interventions. A higher rate of FP may be acceptable for lower risk treatment adjustments, such as optimization of dose of mesalamine, addition of topical therapy, or 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: Variations in fecal calprotectin levels have been documented between different assays tested on the same stool sample.[34] However, the equivalence or interchangeability of calprotectin assays has not been thoroughly evaluated. Five studies directly comparing different assays identified discrepancies ranging from 2.5- to 5-

fold differences between assays when each tested the same stool sample. However, most of this variability occurred at the higher end of the calprotectin range, and was reduced at the lower range of calprotectin values (which are proposed in this guideline).[35,36] One study found only 8%–9% variation within stool samples at lower calprotectin levels, compared with 18%–33% variation at higher calprotectin levels. Several studies also found differences ranging from 13% to 114%, using the same assay repeatedly with different stool samples from the same patient.[36–38] Other studies have found variation in results testing different regions of the same stool sample, ranging from 3% to 31%.[36,38] No similar studies of variability or reproducibility of lactoferrin assays have been published.

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 variations 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: In addition to the accuracy and performance of the biomarker itself as a surrogate for disease activity, there is heterogeneity in the performance of the biomarker for a given patient. Among the included biomarkers, this is best exemplified for CRP.[39] In large genetic studies, the fraction of heritability attributed to CRP has been between 25% and 40%. In a study of 250 healthy US Army recruits, 2 polymorphisms in the CRP gene (−717G>A and +1444C>T) influenced both baseline CRP levels and elevation in response to vigorous exercise.[40] In an inflammatory bowel disease cohort, patients with −717 wild type had higher high-sensitivity CRP concentrations than those with non–wild type.[41] The prevalence of non–wild-type status at the −717 and +1444 locations are estimated to be 10%–15%, with an additional 30%–35% having heterozygosity at these loci. In a study of 199 subjects with active CD, other variants in the CRP gene were also associated with lower degree of CRP elevation in the presence of specific variants.[42] Thus, it is plausible that in patients with these CRP gene variants, the performance of CRP as a biomarker may be less reliable. Therefore, in conjunction with their cross-sectional use, it is important to consider the longitudinal history of CRP elevation within each patient, focusing on their use in those individuals who have previously demonstrated a robust elevation in active inflammation. Similarly, the elevation of fecal calprotectin and lactoferrin may be most accurate in those who have previously demonstrated an elevation. However, it is important to state that studies that have examined the performance of these biomarkers have done so in unselected cohorts, agnostic of inter-individual heterogeneity in potential for elevation of these markers. Thus, the performance of these markers as established in these guidelines remains broadly applicable across populations. The performance may indeed be better in patients where a correlation has been established between biomarkers and endoscopic activity. Thus, one can consider benchmarking the serum and/or fecal biomarkers at the time of endoscopic assessment to determine their correlation in an individual patient.

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 active disease and remission).

Guideline Recommendations

Patients With Ulcerative Colitis in Symptomatic Remission

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

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

Comment: Patients who place high value on avoiding the burden of biomarker testing, over a potentially higher risk of flare or overtreatment, may reasonably choose interval symptom-based monitoring.

Implementation Considerations

- Interval biomarker monitoring may be performed every 6–12 months.
- Fecal biomarkers (fecal calprotectin or fecal lactoferrin) may be optimal for monitoring and may be particularly useful in patients where biomarkers have historically correlated with endoscopic disease activity.
- A biomarker-based monitoring strategy, especially using stool-based tests, however, may be inconvenient and elevated biomarkers in otherwise asymptomatic individuals may lead to high patient anxiety. The diagnostic

performance of these tests in a low pretest probability setting is suboptimal resulting in unacceptably high rates of false results.

- It is important to think about the downstream consequences of testing and associated costs. The optimal management strategy in cases of discrepancy between symptoms and biomarkers is unclear and would generally trigger additional endoscopic testing for confirmation or repeat biomarker testing.

Summary of the Evidence

A biomarker-based monitoring strategy involves routine assessment of symptoms along with noninvasive biomarkers of inflammation to inform ongoing management in patients with UC in symptomatic remission. Supplementary Figure 3 lays out the schematic for the proposed comparison. We did not identify any RCTs that directly compared a biomarker-based monitoring strategy with a symptom-based monitoring strategy. Only 1 RCT examined the impact of mesalamine dosage escalation on reducing fecal calprotectin in patients with quiescent UC.[43] Of 52 patients with mild UC in symptomatic remission with fecal calprotectin >50 μg/g, 26 patients were randomized to increasing mesalamine dosage by 2.4 g/d for 6 weeks vs 26 patients who continued on a stable dosage of mesalamine. In this trial, the primary end point of continued clinical remission with normalization of fecal calprotectin (<50 μg/g) by week 6 was more likely to be achieved in those randomized to escalation of mesalamine (27% vs 4%). However, there were no differences in time to clinical relapse by week 48. This trial did not adequately inform the focused question, given limited duration of intervention (only 6 weeks) and limited information on ongoing monitoring and optimization.

We subsequently examined cohort studies in patients with UC in symptomatic remission in which patients underwent biomarker testing, and long-term outcomes were compared between those with elevated biomarkers and those with normal biomarkers. We posited that if long-term outcomes are significantly different in patients with elevated biomarkers compared with those with normal biomarkers, then an interval biomarker-based monitoring in asymptomatic patients may inform prognosis and long-term management. We identified 17 cohort studies with 1286 patients with UC in symptomatic remission (Supplementary Figure 4). In these studies, fecal calprotectin was the preferred biomarker used for monitoring; 36% of patients were classified as having elevated fecal calprotectin (usually >150 μg/g) and 64% had normal fecal calprotectin. On median follow-up of 1 year, patients with elevated fecal calprotectin were 4.4 times more likely to have disease relapse compared with patients with normal fecal calprotectin (95% CI, 3.48–5.47) with low heterogeneity (I^2 = 24%). With an observed median annual risk of relapse of 15% in patients with UC in symptomatic remission and normal fecal calprotectin in these cohorts, estimated annual risk of relapse in patients with quiescent UC and elevated fecal calprotectin was 64% (Table 5).

Benefits and Harms (Downsides)

Symptom-based monitoring strategy. The potential benefit of a symptom-based monitoring strategy is the convenience of relying only on patient-reported outcomes that can be readily ascertained. However, harms related to a symptom-based monitoring strategy are higher rates of false reassurance and higher risk of disease-related complications (in patients with symptomatic remission but elevated biomarkers who are at higher risk of relapse).

Biomarker-based monitoring strategy. Potential benefits of a biomarker-based monitoring strategy include more accurate prognostication than symptoms alone, to facilitate optimal treatment decisions and lower the risk of disease complications. Potential harms of a biomarker-based monitoring strategy include the costs and inconvenience of sample collection, particularly stool-based tests. In addition, elevated biomarkers in otherwise asymptomatic patients may lead to higher patient anxiety, and with high FP rates in this scenario (see question 2), would require follow-up invasive procedures or repeat biomarker testing.

Table 5. Evidence Profile: Question 1. What Is the Risk of Relapse in Patients With Ulcerative Colitis in Symptomatic Remission With Elevated vs Normal Fecal Calprotectin During Routine Follow-Up?[a]

Outcome/no. of participants (studies)	Relative effect, RR (95% CI)	Anticipated absolute effects (95% CI)			Certainty of evidence
		Normal fecal calprotectin	Elevated fecal calprotectin	Difference	
		Pooled relapse rate, %			
Risk of relapse at 12 mo/ 1286 (17 cohorts)	4.36 (3.48–5.47)	15	65.4 (52.2–82)	50.4 more (37.2 more to 67 more)	⨁⨁⨁◯ MODERATE[b]

RR, relative risk.
[a]Patient or population: patients with UC in symptomatic remission; setting: cohort; exposure: elevated fecal calprotectin (generally >150 μg/g); and comparison: normal fecal calprotectin.
[b]Evidence rated down for risk of bias based on Quality in Prognostic Studies tool and slight variability in fecal calprotectin cutoffs.

Certainty of Evidence

When examining cohort studies comparing long-term outcomes in patients with UC in symptomatic remission with elevated vs normal biomarkers, there was moderate 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 and variability in cutoffs of fecal calprotectin (Table 5). There were limited data on prognostic value of other biomarkers like fecal lactoferrin and serum CRP in patients with asymptomatic UC.

Rationale

Using the GRADE Evidence-to-Decision framework, incorporating the potential benefits and downsides of the 2 strategies and considerations of resource utilization, acceptability, feasibility and equity, the guideline panel conditionally recommended in favor of a biomarker-based monitoring strategy compared with a symptom-based monitoring strategy. Some patients who place higher value on avoiding the burden of biomarker testing over a potentially high risk of flares may reasonably choose interval symptom-based monitoring. Cost-effectiveness analyses have suggested that symptom-based monitoring may be the most cost-effective approach to implement treat-to-target monitoring for patients with UC receiving biologics and small molecule inhibitors.[44]

Several other factors need to be considered when deciding appropriate monitoring strategies. Most of the data on predicting risk of relapse in patients with symptomatic UC was based on fecal calprotectin; as noted below, the diagnostic performance (particularly the sensitivity) of serum CRP is lower and, hence, the prognostic performance of normal CRP in patients with UC in symptomatic remission may not be as informative. The optimal management strategy in case of discrepancy between symptoms and biomarkers is unclear—the diagnostic performance of these tests in a low pretest probability setting is suboptimal, resulting in unacceptably high rates of FP or FN— and would generally trigger additional endoscopic testing for confirmation. The guideline panel felt that in patients with UC in symptomatic remission with elevated biomarkers, repeat biomarker testing may be a reasonable alternative.

Question 2: In patients with UC in symptomatic remission, at what (A) fecal calprotectin, (B) fecal lactoferrin, and (C) serum C-reactive protein cutoff can we accurately rule out active inflammation, obviating routine endoscopic assessment?

Recommendation 2: In patients with UC in symptomatic remission, the AGA suggests using fecal calprotectin <150 μg/g, normal fecal lactoferrin, or normal CRP to rule out active inflammation and avoid routine endoscopic assessment of disease activity. *(Conditional recommendation, very low to low certainty of evidence)*

Implementation Considerations

- In patients who have recently achieved symptomatic remission after treatment adjustment in the preceding 1–3 months, fecal calprotectin <50 μg/g may be preferred over <150 μg/g to detect endoscopic improvement (MES 0 or 1).
- Normal CRP may be less informative to rule out moderate to severe active endoscopic inflammation in patients with UC in symptomatic remission, particularly in patients who have recently achieved symptomatic remission after treatment adjustment. However, if CRP was elevated at time of initial flare, then normalization of CRP may suggest endoscopic improvement (MES 0 or 1).

Recommendation 3: In patients with UC in symptomatic remission but elevated stool or serum markers of inflammation (fecal calprotectin >150 μg/g, elevated fecal lactoferrin, elevated CRP), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, very low certainty of evidence*)

Implementation Consideration

In patients with UC in symptomatic remission but elevated biomarkers of inflammation, repeat measurement of biomarkers (in 3–6 months) may be a reasonable alternative to endoscopic assessment. However, if biomarkers are elevated on repeat evaluation, then endoscopic assessment may be warranted.

Recommendation 4: In patients with UC with mild symptoms, with normal stool or serum markers of inflammation (fecal calprotectin <150 μg/g, normal fecal lactoferrin, normal CRP), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, very low certainty of evidence*)

Implementation Consideration

- In patients with UC with mild symptoms (eg, slight increase in stool frequency and/or infrequent rectal bleeding), it may be reasonable to proceed directly with endoscopic assessment rather than testing biomarkers of inflammation.
- In patients with UC with mild symptoms and normal biomarkers of inflammation who prefer to avoid endoscopic assessment or empiric treatment escalation, repeat measurement of biomarkers (in 3–6 months) may be a reasonable alternative.

Summary of the Evidence

1. In patients with UC in symptomatic remission (no rectal bleeding, normal or near-normal stool frequency), fecal

calprotectin <150 ± 50 μg/g and normal fecal lactoferrin reliably rules out active inflammation, obviating endoscopic assessment (*low certainty of evidence*); normal serum CRP may rule out active inflammation (*very low certainty of evidence*).

2. In patients with UC in symptomatic remission, elevated fecal calprotectin >150 ± 50 μg/g, elevated fecal lactoferrin, or elevated CRP may not indicate active inflammation (*very low certainty of evidence*).

3. In patients with UC who have mild symptoms (infrequent rectal bleeding and/or increased stool frequency), fecal calprotectin <150 ± 50 μg/g, normal fecal lactoferrin, or normal CRP cannot rule out active inflammation (*very low certainty of evidence*).

Diagnostic Performance of Fecal Calprotectin

The evidence synthesis team decided *a priori* to examine the following 3 diagnostic cutoffs for fecal calprotectin most frequently studied and used in clinical practice: 50 μg/g, 150 μg/g, and 250 μg/g. To account for variability in reported cutoffs in studies, we allowed for values 50 μg/g above and below the cutoff. In the diagnostic testing spectrum, lower cutoffs are more sensitive and higher cutoffs are more specific. We conducted a systematic review to identify cross-sectional and cohort studies in patients with established UC, which reported the diagnostic accuracy of fecal calprotectin for detecting moderate to severe endoscopic inflammation (MES 2 or 3). From these studies, 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. Using this approach, the sensitivity and specificity of fecal calprotectin cutoff of 50 ± 50 μg/g was 78% (95% CI, 66%–86%) and 57% (95% CI, 40%–72%), respectively, based on 11 cohorts; corresponding sensitivity and specificity of 150 ± 50 μg/g cutoff (12 cohorts) was 71% (95% CI, 62%–78%) and 69% (95% CI, 62%–75%), respectively, and of 250 ± 50 μg/g cutoff (9 cohorts) was 67% (95% CI, 53%–78%) and 73% (95% CI, 65%–80%), respectively (Supplementary Figure 5).

Low pretest probability (asymptomatic patients with rectal bleeding score 0 and stool frequency score 0 or 1, with 15% prevalence of moderate to severe inflammation). In applying these cutoffs to a low pretest probability scenario, approximately 3.3%, 4.3%, and 5.5% patients (FN rate) with fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g, respectively, may be misclassified as having endoscopic improvement (MES 0 or 1) when they actually have moderate to severe endoscopic activity (MES 2 or 3) (Table 6). 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 (36.6%, 26.4%, and 23%, respectively), that is, a significant proportion of patients who have endoscopic improvement (MES 0 or 1) may be incorrectly classified as having moderate to severe endoscopic activity.

Intermediate pretest probability (patients with mild symptoms of ulcerative colitis, such as infrequent rectal bleeding [rectal bleeding score 1], or increased stool frequency [stool frequency score 2 or 3], with 50% prevalence of moderate to severe inflammation). In an intermediate pretest probability scenario, approximately 11%, 14.5%, and 18.5% patients (FN rate) with fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g, respectively, may be misclassified as having endoscopic improvement (MES 0 or 1) when they actually have moderate to severe endoscopic activity (MES 2 or 3) (Table 6). 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 (21.5%, 15.5%, and 13.5%, respectively), that is, a significant proportion of patients who have endoscopic improvement may be incorrectly classified as having moderate to severe endoscopic activity.

Diagnostic Performance of Fecal Lactoferrin

The evidence base for fecal lactoferrin was more limited. We identified 9 studies reporting the diagnostic accuracy of fecal lactoferrin for detecting moderate to severe endoscopic inflammation (defined as MES 2 or 3 in 4 studies, and MES 1, 2, or 3 in 5 studies). Studies reported performance of only a single lactoferrin cutoff within a range of 7.25–10 μg/g; the commercial assay reports lactoferrin as positive (elevated) or negative, corresponding to a cutoff of 7.25 μg/g. At this cutoff, the sensitivity and specificity of fecal lactoferrin for detecting endoscopic inflammation was 83% (95% CI, 72%–90%) and 75% (95% CI, 59%–87%), respectively (Supplementary Figure 6).

Low pretest probability (asymptomatic patients with rectal bleeding score 0 and stool frequency score 0 or 1, with 15% prevalence of moderate to severe inflammation). In applying this cutoff to a low pretest probability scenario, approximately 2.6% patients (FN rate) with normal fecal lactoferrin (<7.25 μg/g) may be misclassified as having endoscopic improvement (MES 0 or 1) when they actually have moderate to severe endoscopic activity (MES 2 or 3) (Table 7). In contrast, elevated fecal lactoferrin (>7.25 μg/g), in this low pretest probability scenario, had significantly high rates of being FP (21.2%), that is, 21.2% patients who have endoscopic improvement may be incorrectly classified as having moderate to severe endoscopic activity.

Intermediate pretest probability (patients with mild symptoms of ulcerative colitis, such as infrequent rectal bleeding [rectal bleeding score 1], or increased stool frequency [stool frequency score 2 or 3], with 50% prevalence of moderate to severe inflammation). In an intermediate pretest probability scenario, approximately 8.5% patients (FN rate) with fecal lactoferrin <7.25 μg/g may be misclassified as having endoscopic improvement when they actually have moderate to severe endoscopic activity (Table 7). In contrast, elevated fecal lactoferrin (>7.25 μg/g), in this intermediate pretest probability scenario, had significantly high rates of being FP (12.5%).

Table 6. Evidence Profile: Question 2. (A) Fecal Calprotectin: In Patients With Ulcerative Colitis in Symptomatic Remission, How Accurate Is Fecal Calprotectin Cutoff of <50 μg/g vs <150 μg/g vs <250 μg/g for Ruling Out Moderate to Severe Endoscopically Active Disease (Mayo Endoscopic Score 2 or 3), Obviating the Need for Routine Endoscopic Assessment?[a]

	No. of results per 1000 patients tested (95% CI)						
	Low likelihood (prevalence 15%)			Intermediate likelihood (prevalence 50%)			
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)	117 (99–129)	107 (93–117)	95 (80–117)	390 (330–430)	355 (310–390	315 (265–390)	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)	33 (21–51)	43 (33–57)	55 (33–70)	110 (70–170)	145 (110–190)	185 (110–235)	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]	LOW[b,c]	VERY LOW[b,d]	VERY LOW[b,d]	VERY LOW[b,d]	VERY LOW[b,d]	—
TNs (patients correctly diagnosed as being in endoscopic remission or having only mild endoscopic activity)	484 (340–612)	586 (527–638	620 (553–680)	285 (200–360)	345 (310–375)	365 (325–400)	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
	Low likelihood (prevalence 15%)			Intermediate likelihood (prevalence 50%)			
	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)	366 (238–510)	264 (212–323)	230 (170–297)	215 (140–300)	155 (125–190)	135 (100–175)	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: Patients with UC in symptomatic remission; low pretest probability/likelihood of having moderate to severe endoscopically active disease (no rectal bleeding [RBS 0], and normal to mild increase in stool frequency [SFS 0 or 1], under routine maintenance therapy, or having recently achieved symptomatic remission after treatment adjustment) with estimated prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 15%; intermediate pretest probability/likelihood of having moderate to severe endoscopically active disease (patients with mild symptoms of UC, such as infrequent rectal bleeding [RBS 0 or 1] and/or increased stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 50%.
Pooled sensitivity/specificity fecal calprotectin with cutoff <50 μg/g: sensitivity, 78% (95% CI, 66%–86%); specificity, 57% (95% CI, 40%–72%); 11 studies.
Pooled sensitivity/specificity fecal calprotectin with cutoff <150 μg/g: sensitivity, 71% (95% CI, 62%–78%); specificity, 69% (95% CI, 62%–75%); 12 studies.
Pooled sensitivity/specificity fecal calprotectin with cutoff <250 μg/g: sensitivity, 67% (95% CI, 53%–78%); specificity, 73% (95% CI, 65%–80%); 9 studies.
Reference test: lower endoscopy.
[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 is higher than maximal tolerable FN threshold.
[e]Very serious imprecision because point estimate is higher than maximal tolerable FP threshold.

Table 7. Evidence Profile: Question 2. (B) Fecal Lactoferrin. In Patients With Ulcerative Colitis in Symptomatic Remission, How Accurate Is Negative Fecal Lactoferrin for Ruling Out Moderate to Severe Endoscopically Active Disease (Mayo Endoscopic Score 2 or 3), Obviating the Need for Routine Endoscopic Assessment?[a]

	No. of results per 1000 patients tested (95% CI)		
	Low-likelihood (prevalence 15%)	Intermediate-likelihood (prevalence 50%)	
Test result	Negative lactoferrin	Negative lactoferrin	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	124 (108–135)	415 (360–450)	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)	26 (15–42)	85 (50–140)	FNs would be falsely reassured, and may be at higher risk of disease complications/flare due to undertreatment.
GRADE certainty of evidence	LOW[b]	VERY LOW[b,c]	—
TNs (patients correctly diagnosed as being in endoscopic remission or having only mild endoscopic activity)	638 (501–739)	375 (295–435)	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)	212 (111–349)	125 (65–205)	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]	VERY LOW[b,d]	—

[a]Population/setting: patients with UC in symptomatic remission; low pretest probability/likelihood of having moderate to severe endoscopically active disease (no rectal bleeding [RBS 0], and normal to mild increase in stool frequency [SFS 0 or 1], under routine maintenance therapy, or having recently achieved symptomatic remission after treatment adjustment) with estimated prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 15%; intermediate pretest probability/likelihood of having moderate to severe endoscopically active disease (patients with mild symptoms of UC, such as infrequent rectal bleeding [RBS 0 or 1] and/or increased stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active disease (MES 2or 3) of 50%.
Pooled sensitivity/specificity fecal lactoferrin with cutoff $<$7.25–10 μg/g: sensitivity, 83% (95% CI, 72%–90%); specificity, 75% (95% CI, 59%–87%); 9 studies.
Reference test: lower endoscopy.
[b]Very serious inconsistency, due to selective reporting of cutoffs in studies optimized for best performance and high heterogeneity for summary sensitivity/specificity.
[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.

Diagnostic Performance of Serum C-Reactive Protein

We identified 15 studies reporting the diagnostic accuracy of serum CRP for detecting moderate to severe endoscopic inflammation. Studies reported performance of only a single CRP cutoff with a range of 1.2–7.3 mg/L. Summary sensitivity and specificity of elevated CRP for detecting endoscopic inflammation was 63% (95% CI, 50%–75%) and 77% (95% CI, 67%–84%), respectively (Supplementary Figure 7).

Low pretest probability (asymptomatic patients with rectal bleeding score 0 and stool frequency score 0 or 1, with 15% prevalence of moderate to severe inflammation). In applying this cutoff (elevated CRP, generally >5 mg/L) to a low pretest probability scenario, approximately 5.5% patients (FN rate) with normal CRP (<5 mg/L) may be misclassified as having endoscopic improvement when they actually have moderate to severe endoscopic activity (Table 8). In contrast, elevated CRP (>5 mg/L), in this low pretest probability scenario, had significantly high rates of being FP (19.5%), that is, 19.5% patients who have endoscopic improvement may be incorrectly classified as having moderate to severe endoscopic activity.

Intermediate pretest probability (patients with mild symptoms of ulcerative colitis, such as infrequent rectal bleeding [rectal bleeding score 1], or increased stool frequency [stool frequency score 2 or 3], with 50% prevalence of moderate to severe inflammation). In an intermediate pretest probability scenario, approximately 18.5% of patients (FN rate) with normal CRP (<5 mg/L) may be misclassified as having endoscopic improvement when they actually have moderate to severe endoscopic remission (MES 2 or 3) (Table 8). In contrast, elevated CRP (>5 mg/L), in this intermediate pretest probability scenario, had significantly high rates of being FP (11.5%).

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 moderate to severe 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 fecal calprotectin cutoffs of <50 μg/g and <150 μg/g to rule out moderate to severe endoscopic inflammation in a low pretest probability setting, and very low certainty of evidence supporting the use of fecal calprotectin cutoffs of <250 μg/g in this scenario (Table 6). Evidence was rated down for inconsistency due to selective inclusion of studies reporting specific cutoffs and high heterogeneity for summary sensitivity and specificity, and for imprecision because 95% CI of the maximal tolerable FN rate was 5%; evidence for the cutoff of <250 μg/g was further rated down for very serious imprecision because the point estimate is higher than the maximal tolerable FN rate.

In contrast, in the intermediate probability scenario, there was very low certainty of evidence supporting the use of any proposed fecal calprotectin cutoff to rule out moderate to severe endoscopic inflammation due to unacceptably high rates of FN (very serious imprecision, because the point estimate crossed the FN threshold of 5%) and selective inclusion of studies and heterogeneity in summary sensitivity and specificity (inconsistency) (Table 6). Similarly, in the low and intermediate probability scenario, there was very low certainty of evidence supporting the use of any proposed cutoff of elevated fecal calprotectin to rule in moderate to severe endoscopic inflammation due to unacceptably high rates of FP (very serious imprecision) and inconsistency.

Fecal lactoferrin. There was low certainty of evidence supporting the use of normal fecal lactoferrin to rule out moderate to severe endoscopic inflammation in a low pretest probability setting (Table 7). Evidence was rated down for very serious inconsistency, due to selective reporting of cutoffs in studies optimized for best performance and high heterogeneity for summary sensitivity and specificity. In contrast, in the intermediate probability scenario, there was very low certainty of evidence supporting the use of normal fecal lactoferrin to rule out moderate to severe endoscopic inflammation, due to unacceptably high rates of FN (very serious imprecision) and inconsistency. Similarly, in the low and intermediate probability scenario, there was very low certainty of evidence supporting the use of elevated fecal lactoferrin to rule in moderate to severe endoscopic inflammation, due to unacceptably high rates of FP (very serious imprecision) and inconsistency (Table 7).

Serum C-reactive protein. There was very low certainty of evidence supporting the use of normal CRP to rule out moderate to severe endoscopic inflammation in a low pretest probability setting (Table 8). 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 very serious imprecision because the point estimate is higher than the maximal tolerable FN rate. Similarly, in the intermediate probability scenario, there was very low certainty of evidence supporting the use of normal serum CRP to rule out moderate to severe endoscopic inflammation, due to unacceptably high rates of FN (very serious imprecision) and inconsistency (Table 8). In the low and intermediate probability scenario, there was very low certainty of evidence supporting the use of elevated CRP to rule in moderate to severe endoscopic inflammation, due to unacceptably high rates of FP (very serious imprecision) and inconsistency.

Rationale

In using noninvasive biomarkers as a triage strategy to determine need for endoscopy and ongoing management, health care providers and patients need to be aware of test performance and the downstream consequences of potential FN and FP rates. The guideline panel and evidence synthesis team determined *a priori* a maximal tolerable FN threshold of 5% for patients with UC in symptomatic remission. However, the team deemed that there may be circumstances when patients and providers may be willing to accept higher rates of FN, depending on risk of downstream consequences, particularly the nature of treatment adjustment, and emphasize the importance of shared decision making.

Table 8. Evidence Profile: Question 2. (C) Serum C-Reactive Protein: In Patients With Ulcerative Colitis in Symptomatic Remission, How Accurate Is Normal Serum C-Reactive Protein for Ruling Out Moderate to Severe Endoscopically Active Disease (Mayo Endoscopic Score 2 or 3), Obviating the Need for Routine Endoscopic Assessment?[a]

	No. of results per 1000 patients tested (95% CI)		
	Low-likelihood (prevalence 15%)	Intermediate-likelihood (prevalence 50%)	
Test result	Normal CRP	Normal CRP	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	95 (75–112)	315 (250–375)	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)	55 (38–75)	185 (125–250)	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)	655 (570–714)	385 (335–420)	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)	195 (136–280)	115 (80–165)	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]	VERY LOW[b,d]	—

[a]Population/setting: patients with UC in symptomatic remission; low pretest probability/likelihood of having moderate to severe endoscopically active disease (no rectal bleeding [RBS 0], and normal to mild increase in stool frequency [SFS 0 or 1], under routine maintenance therapy, or having recently achieved symptomatic remission after treatment adjustment) with estimated prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 15%; intermediate pretest probability/likelihood of having moderate to severe endoscopically active disease (patients with mild symptoms of UC, such as infrequent rectal bleeding [RBS 0 or 1] and/or increased stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 50%.
Pooled sensitivity/specificity CRP with cutoff <5 mg/L: sensitivity, 63% (95% CI, 50%–75%); specificity, 77% (95% CI, 67%–84%); 15 studies.
Reference test: lower endoscopy.
[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.

For ease of implementation in clinical practice, the guideline panel felt 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 multiple different cutoffs for different scenarios. There may be circumstances, such as patients who may have recently achieved symptomatic remission after treatment adjustment in the preceding 1–3 months, when a lower fecal calprotectin <50 μg/g may be more accurate than <150 μg/g to rule out the presence of moderate to severe active inflammation. It is important to note that in children 2 years or younger, a higher threshold for fecal calprotectin may be needed due to a wider range of normal calprotectin in young children.[45,46]

The guideline panel did not compare the performance of different noninvasive biomarkers, due to variable cutoffs for each test. Stool-based tests may be more sensitive for intestinal inflammation compared with serum CRP; however, CRP has the convenience of being a blood test. The panel noted that fecal calprotectin has been well-studied and was able to study the performance of different cutoffs to adequately ascertain performance. In contrast, fecal lactoferrin had a limited evidence base with limited studies on different cutoffs. Similarly, CRP had a limited evidence base despite being very commonly measured in clinical practice. There may be circumstances when the performance of CRP may be suboptimal. For example, in patients who have recently achieved symptomatic remission after treatment adjustment, normal CRP may be less informative, exceeding the FN threshold. However, if the CRP was elevated at time of initial flare, then normalization of CRP may suggest endoscopic improvement. The panel did not study proprietary tests that are not widely available or indicated for use in UC.

Patients With Symptomatically Active Ulcerative Colitis

Question 3: In patients with symptomatically active UC, 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 UC, the AGA suggests an evaluation strategy that combines biomarkers and symptoms, rather than symptoms alone, to inform treatment adjustments. (*Conditional recommendation, low certainty of evidence*)

Comment: Patients, particularly those with severe symptoms, who place a high value on avoiding the burden of biomarker testing, over a potentially higher risk of inappropriate overtreatment, may reasonably choose symptom-based evaluation for treatment decisions.

Summary of the Evidence

A biomarker-based evaluation strategy involves checking noninvasive biomarkers of inflammation in patients with symptomatically active UC to inform ongoing management; in contrast, symptom-based evaluation would involve treatment decisions being driven based solely on symptoms. We did not identify any RCTs that directly compared a biomarker-based evaluation strategy with symptom-based evaluation for patients with symptomatically active UC. Recognizing that presence of moderate to severe endoscopic inflammation, in conjunction with symptoms, is a key trigger for treatment decisions, we used ability of symptoms alone vs symptoms plus biomarkers to detect presence of moderate to severe inflammation on endoscopy as a surrogate outcome to inform decision making and improve patient outcomes. In a prior pooled analysis of 6 clinical trials of biologic agents and tofacitinib in 2586 patients with moderate to severely active UC, authors examined the cross-sectional prevalence of moderate to severe endoscopic inflammation (based on MES 2 or 3) in patients with varying combinations of cardinal symptoms of UC (RBS and SFS components of MES).[21] In this analysis, 85%–90% patients with RBS 2 or 3 and SFS 2 or 3 had moderate to severe inflammation on endoscopy (Supplementary Table 1). This suggests an FP rate of 10%–15%, that is, 10%–15% patients with typical symptoms suggestive of active UC may be in endoscopic remission or have only mildly active disease, such that relying on symptoms alone may lead to potentially unnecessary treatment adjustments (such as adding corticosteroids, escalating or switching therapies). As shown in subsequent analyses (see question 4), in patients with typical symptoms suggestive of active UC (RBS 2 or 3 and SFS 2 or 3), presence of elevated biomarkers of inflammation decreases the FP rate to $<5\%$; that is, $<5\%$ patients with symptoms and elevated biomarkers will actually have only mild inflammation or be in endoscopic remission, resulting in acceptably low rates of unnecessary treatment adjustments.

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 and faster decision making. However, harms related to relying on symptoms only are higher rates of inappropriate treatment adjustments or overtreatment and treatment-related complications (in case of 10%–15% of patients with symptoms suggestive of UC but who may be in endoscopic remission or have only mildly active disease).

Biomarker-based evaluation strategy. Potential benefits of a biomarker-based evaluation strategy is more accurate prognostication than symptoms alone to facilitate optimal treatment decisions and avoid 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 happen due to the extra step of test completion.

Certainty of Evidence

In the absence of randomized trials, we relied on cross-sectional studies, with indirect comparisons and surrogate outcome (presence of moderate to severe endoscopic inflammation), with somewhat imprecise estimates with a variety of biomarkers. Hence, there was low confidence in effect estimates supporting a biomarker-based evaluation strategy over symptom-based evaluation in patients with UC with active symptoms.

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 a symptom-based evaluation alone in patients with symptomatically active UC. The panel recognized that adding an extra step of biomarker testing in patients with symptomatically active UC may potentially delay treatment 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, and who place high value in avoiding burden of biomarker testing, may reasonably choose symptom-based evaluation for treatment decisions, acknowledging a 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.

Optimal management strategy in case of discrepancy between symptoms and biomarkers is unclear. In patients with typical symptoms suggestive of UC, normal biomarkers may not exclude lack of moderate to severe inflammation, and endoscopic assessment may be a preferred approach. However, in a subset of patients, noninvasive biomarkers do not correlate with endoscopic inflammation. In this setting, shared decision making on empiric treatment adjustment with a 10%–15% FP rate of symptoms alone may be acceptable, particularly when access to endoscopy is limited (which may lead to delay in treatment initiation) and treatment adjustments being considered are low risk.

Question 4: In patients with symptomatically active UC, at what (A) fecal calprotectin, (B) fecal lactoferrin, and (C) serum C-reactive protein cutoff can we accurately diagnose active inflammation, obviating routine endoscopic assessment?

Recommendation 6: In patients with UC with moderate to severe symptoms suggestive of flare, the AGA suggests using fecal calprotectin >150 μg/g, elevated fecal lactoferrin, or elevated CRP to rule in active inflammation and inform treatment adjustment and avoid endoscopic assessment solely for establishing presence of active disease. (*Conditional recommendation, very low to moderate certainty of evidence*)

Comment: Patients who place greater value in confirming inflammation, particularly when making significant treatment adjustments (such as starting or switching immunosuppressive therapies) and lesser value on the inconvenience, cost, or risk of endoscopy, may choose to pursue endoscopic evaluation before treatment adjustment.

Recommendation 7: In patients with UC with mild symptoms, with elevated stool or serum markers of inflammation (fecal calprotectin >150 μg/g, elevated fecal lactoferrin, or elevated CRP), the AGA suggests endoscopic assessment of disease activity rather than empiric treatment adjustment. (*Conditional recommendation, very low certainty of evidence*)

Implementation Consideration

In patients with UC who underwent recent adjustment of treatment in response to moderate to severe symptomatic flare, and now have mild residual symptoms, elevated stool or serum markers of inflammation may be used to inform treatment adjustments (such as dose adjustments of therapy).

Summary of the Evidence

1. In patients with UC with moderate to severe symptoms suggestive of flare (frequent rectal bleeding, significantly increased stool frequency), fecal calprotectin >150 μg/g, elevated fecal lactoferrin, and elevated CRP reliably suggest moderate to severe endoscopic inflammation, obviating routine need for endoscopic assessment (*very low to moderate certainty of evidence*).
2. In patients with UC who have mild symptoms (infrequent rectal bleeding and/or increased stool frequency), fecal calprotectin >150 μg/g, elevated fecal lactoferrin, and elevated CRP may not suggest moderate to severe endoscopic inflammation (*very low certainty of evidence*).
3. In patients with UC with moderate to severe symptoms suggestive of flare (frequent rectal bleeding, significantly increased stool frequency), fecal calprotectin <150 μg/g, normal fecal lactoferrin or normal CRP may not suggest lack of inflammation (*very low certainty of evidence*).

Diagnostic Performance of Fecal Calprotectin

Summary sensitivity and specificity of fecal calprotectin for detecting moderate to severe endoscopic inflammation has been reported in question 2.

High pretest probability scenario (patients with typical symptoms of ulcerative colitis flare with frequent rectal bleeding [rectal bleeding score 2 or 3] and significant increase in stool frequency [stool frequency score 2 or 3], with 85% prevalence of moderate to severe inflammation). In applying these cutoffs in high pretest probability scenarios, approximately 6.4%, 4.6%, and 4.0% patients (FP rate) with fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g, respectively, may be misclassified as having moderate to severe endoscopic activity (MES 2 or 3) when they actually have endoscopic improvement (MES 0 or 1) (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 (18.7%, 24.7%, and 31.4%, respectively), that is, a significant proportion of symptomatic patients who have moderate to severe endoscopic activity may be incorrectly classified as having endoscopic improvement.

Intermediate pretest probability scenario (patients with mild symptoms of ulcerative colitis, such as infrequent rectal bleeding [rectal bleeding score 1], or increased stool frequency [stool frequency score 2 or 3],

with 50% prevalence of moderate to severe inflammation). In an intermediate pretest probability scenario, approximately 21.5%, 15.5%, and 13.5% patients (FP rate) with fecal calprotectin >50 μg/g, >150 μg/g, and >250 μg/g, respectively, may be misclassified as having moderate to severe endoscopic activity (MES 2 or 3) when they actually have endoscopic improvement (MES 0 or 1) (Table 9). In contrast, fecal calprotectin <50 μg/g, <150 μg/g, and <250 μg/g in this high pretest probability scenario, had high rates of being FN (11.0%, 14.5%, and 18.5%, respectively), that is, a significant proportion of symptomatic patients who have moderate to severe endoscopic activity may be incorrectly classified as having endoscopic improvement.

Diagnostic Performance of Fecal Lactoferrin

Summary sensitivity and specificity of fecal lactoferrin for detecting moderate to severe endoscopic inflammation has been reported above in question 2.

High pretest probability scenario (patients with typical symptoms of ulcerative colitis flare with rectal bleeding score 2 or 3 and stool frequency score 2 or 3, with 85% prevalence of moderate to severe inflammation). In applying this cutoff to a high pretest probability scenario, approximately 3.7% of patients (FP rate) with elevated fecal lactoferrin (>7.25 μg/g) may be misclassified as having moderate to severe endoscopic activity (MES 2 or 3) (Table 10). In contrast, normal fecal lactoferrin (<7.25 μg/g) had significantly high rates of being FN (14.5%), that is, 14.5% of patients who have moderate to severe endoscopic activity are classified as being in endoscopic improvement.

Intermediate pretest probability scenario (patients with mild symptoms of ulcerative colitis, rectal bleeding score 1, or stool frequency score 2 or 3, with 50% prevalence of moderate to severe inflammation). In an intermediate pretest probability scenario, FP rate of elevated fecal lactoferrin was 12.5% and FN rate of normal fecal lactoferrin was 8.5% (Table 10).

Diagnostic Performance of Serum C-Reactive Protein

Summary sensitivity and specificity of serum CRP for detecting moderate to severe endoscopic inflammation has been reported in question 2.

High pretest probability scenario (patients with typical symptoms of ulcerative colitis flare with rectal bleeding score 2 or 3 and stool frequency score 2 or 3, with 85% prevalence of moderate to severe inflammation). In applying this cutoff (elevated CRP, generally >5 mg/L) to a high pretest probability scenario, approximately 3.4% patients (FP rate) with elevated CRP (>5 mg/L) may be misclassified as having moderate to severe endoscopic activity (MES 2 or 3) (Table 11). In contrast, normal CRP (<5 mg/L) had significantly high rates of being FN (31.4%).

Intermediate pretest probability scenario (patients with mild symptoms of ulcerative colitis, rectal bleeding score 1, or stool frequency score 2 or 3, with 50% prevalence of moderate to severe inflammation). In an intermediate pretest probability scenario, FP rate of elevated CRP was 11.5% and FN rate of normal CRP was 18.5% (Table 11).

Certainty of Evidence

Although there was 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 moderate to severe 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 fecal calprotectin cutoffs of >150 μg/g and >250 μg/g to rule in moderate to severe endoscopic inflammation in a high pretest probability setting (evidence rated down for inconsistency and imprecision), and very low certainty of evidence supporting the use of fecal calprotectin cutoffs of >50 μg/g in this scenario (evidence rated down for inconsistency and very serious imprecision) (Table 9).

In contrast, in the intermediate probability scenario, there was very low certainty of evidence supporting the use of any proposed fecal calprotectin cutoff to rule in moderate to severe endoscopic inflammation, due to unacceptably high rates of FP (very serious imprecision) and selective inclusion of studies and heterogeneity in summary sensitivity and specificity (inconsistency) (Table 9). Similarly, in the high and intermediate probability scenario, there was very low certainty of evidence supporting the use of any proposed cutoff of normal fecal calprotectin to rule out moderate to severe endoscopic inflammation, due to unacceptably high rates of FN (very serious imprecision) and inconsistency.

Fecal lactoferrin. There was very low certainty of evidence supporting the use of elevated fecal lactoferrin to rule in moderate to severe endoscopic inflammation in a high pretest probability setting (Table 10). Evidence was rated down for very serious inconsistency due to selective reporting of cutoffs in studies optimized for best performance and high heterogeneity for summary sensitivity and specificity, and for imprecision (upper limit of 95% CI crossing 5% FP threshold). In the intermediate probability scenario, there was very low certainty of evidence supporting the use of elevated fecal lactoferrin to rule in moderate to severe endoscopic inflammation, due to unacceptably high rates of FP (very serious imprecision) and very serious inconsistency. Similarly, in the high and intermediate probability scenario, there was very low certainty of evidence supporting the use of normal fecal lactoferrin to rule out moderate to severe endoscopic inflammation, due to unacceptably high rates of FN (very serious imprecision) and very serious inconsistency (Table 10).

Serum C-reactive protein. There was moderate certainty of evidence supporting the use of elevated CRP to rule in moderate to severe endoscopic inflammation in a high

Table 9. Evidence Profile: Question 3. (A) Fecal Calprotectin: In Patients With Symptomatically Active Ulcerative Colitis, How Accurate Is Fecal Calprotectin Cutoff of >50 μg/g vs >150 μg/g vs >250 μg/g for Ruling in Moderate to Severe Endoscopically Active Disease (Mayo Endoscopic Score 2 or 3), Obviating the Need for Routine Endoscopic Assessment?[a]

	No. of results per 1000 patients tested (95% CI)						
	Intermediate-likelihood (prevalence 50%)			High-likelihood (prevalence 85%)			
Test result	fCal >50	fCal >150	fCal>250	fCal >50	fCal >150	fCal >250	Comments
TPs (patients with moderate to severe endoscopically active disease)	390 (330–430)	355 (310–390	315 (265–390)	663 (561–731)	603 (527–663)	536 (451–663)	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 classified as being in endoscopic remission or having mildly active disease)	110 (70–170)	145 (110–190)	185 (110–235)	187 (119–289)	247 (187–323)	314 (187–399)	FNs may be falsely reassured, undertreated, or mistreated (as not having UC flare), potentially leading to increased disease related complications and morbidity.
GRADE certainty of evidence	VERY LOW[b,c]	VERY LOW[b,c]	VERY LOW[b,c]	VERY LOW[b,c]	VERY LOW[b,c]	VERY LOW[b,c]	—
TNs (patients in endoscopic remission or having mildly active disease)	285 (200–360)	345 (310–375)	365(325–400)	86 (60–108)	104 (93–113)	110 (98–120)	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 classified as having moderate to severe endoscopically active disease)	215 (140–300)	155 (125–190)	135 (100–175)	64 (42–90)	46 (37–57)	40 (30–52)	FPs may undergo unnecessary treatment adjustment and have treatment-related complications.
GRADE certainty of evidence	VERY LOW[b,d]	VERY LOW[b,d]	VERY LOW[b,d]	VERY LOW[b,c]	LOW[b,e]	LOW[b,e]	—

fCal, fecal calprotectin.
[a]Population/setting: patients with symptomatically active UC; intermediate pretest probability of having moderate to severe endoscopically active disease (patients with mild symptoms of UC, such as infrequent rectal bleeding [RBS 0 or 1] and/or increased stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 50%; high pretest probability of having moderate to severe endoscopically active disease (patients with typical symptoms of UC flare with frequent rectal bleeding [RBS 2 or 3] and significant increase in stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active diserase (MES 2 or 3) of 85%.
Pooled sensitivity/specificity fecal calprotectin with cutoff >50 μg/g: sensitivity, 78% (95% CI, 66%–86%); specificity, 57% (95% CI, 40%–72%); 11 studies.
Pooled sensitivity/specificity fecal calprotectin with cutoff >150 μg/g: sensitivity, 71% (95% CI, 62%–78%); specificity, 69% (95% CI, 62%–75%); 12 studies.
Pooled sensitivity/specificity fecal calprotectin with cutoff >250 μg/g: sensitivity, 67% (95% CI, 53%–78%); specificity, 73% (95% CI, 65%–80%); 9 studies.
Reference test: lower endoscopy.
[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%.

Table 10. Evidence Profile: Question 3. (B) Fecal Lactoferrin: In Patients With Symptomatically Active Ulcerative Colitis, How Accurate Is Positive Fecal Lactoferrin for Ruling in Moderate to Severe Endoscopically Active Disease (Mayo Endoscopic Score 2 or 3), Obviating the Need For Routine Endoscopic Assessment?[a]

	No. of results per 1000 patients tested (95% CI)		
	Intermediate-likelihood (prevalence 50%)	High-likelihood (prevalence 85%)	
Test result	Positive lactoferrin	Positive lactoferrin	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	415 (360–450)	705 (612–765)	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)	85 (50–140)	145 (85–238)	FNs may be falsely reassured, undertreated, or mistreated (as not having UC 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)	375 (295–435)	113 (89–131)	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)	125 (65–205)	37 (19–61)	FPs may undergo unnecessary treatment adjustment and have treatment-related complications.
GRADE certainty of evidence	VERY LOW[b,d]	VERY LOW[b,e]	—

[a]Population/setting: patients with symptomatically active UC; intermediate pretest probability of having moderate to severe endoscopically active disease (patients with mild symptoms of UC, such as infrequent rectal bleeding [RBS 0 or 1] and/or increased stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 50%; high pretest probability of having moderate to severe endoscopically active disease (patients with typical symptoms of UC flare with frequent rectal bleeding [RBS 2 or 3] and significant increase in stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 85%.
Pooled sensitivity/specificity fecal lactoferrin with cutoff $<$7.25-10 μg/g: sensitivity 83% (95% CI, 72%–90%); specificity 75% (95% CI, 59%–87%); 9 studies.
Reference test: lower endoscopy.
[b]Very serious inconsistency heterogeneity, due to selective reporting of cutoffs in studies optimized for best performance and high heterogeneity for summary sensitivity/specificity.
[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%.

Table 11. Evidence Profile: Question 3. (C) Serum C-Reactive Protein: In Patients With Symptomatically Active Ulcerative Colitis, How Accurate Is Elevated Serum C-Reactive Protein for Ruling in Moderate to Severe Endoscopically Active Disease (Mayo Endoscopic Score 2 or 3), Obviating the Need for Routine Endoscopic Assessment?[a]

	No. of results per 1000 patients tested (95% CI)		
	Intermediate-likelihood (prevalence 50%)	High-likelihood (prevalence 85%)	
Test result	Elevated CRP	Elevated CRP	Comments
TPs (patients correctly diagnosed as having moderate to severe endoscopic activity)	315 (250–375)	536 (425–638)	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)	185 (125–250)	314 (212–425)	FNs may be falsely reassured, undertreated, or mistreated (as not having UC 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)	385 (335–420)	116 (101–126)	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)	115 (80–165)	34 (24–49)	FPs may undergo unnecessary treatment adjustment and have treatment-related complications.
GRADE certainty of evidence	VERY LOW[b,d]	MODERATE[b]	—

[a]Population/setting: patients with symptomatically active UC; intermediate pretest probability of having moderate to severe endoscopically active disease (patients with mild symptoms of UC, such as infrequent rectal bleeding [RBS 0 or 1] and/or increased stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 50%; high pretest probability of having moderate to severe endoscopically active disease (patients with typical symptoms of UC flare with frequent rectal bleeding [RBS 2 or 3] and significant increase in stool frequency [SFS 2 or 3]) with observed prevalence of moderate to severe endoscopically active disease (MES 2 or 3) of 85%.
Pooled sensitivity/specificity CRP with cutoff <5 mg/L: sensitivity, 63% (95% CI, 50%–75%); specificity, 77% (95% CI, 67%–84%); 15 studies.
Reference test: lower endoscopy.
[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.

pretest probability setting (Table 11). Evidence was rated down for inconsistency due to selective reporting of cutoffs in studies optimized for best performance and high heterogeneity for summary sensitivity or specificity. In contrast, in the intermediate probability scenario, there was very low certainty of evidence supporting the use of elevated serum CRP to rule in moderate to severe endoscopic inflammation (inconsistency, very serious imprecision). Similarly, in the high and intermediate probability scenario, there was very low certainty of evidence supporting the use of normal CRP to rule out moderate to severe endoscopic inflammation, due to unacceptably high rates of FN (very serious imprecision) and inconsistency (Table 11).

Rationale

The guideline panel and evidence synthesis team determined *a priori* the maximal tolerable FP thresholds at 5% for patients with symptomatically active UC. However, the guideline panel deemed there may be circumstances when patients and providers may be willing to accept higher rates of FP, depending on risk of downstream consequences, particularly the nature of treatment adjustment, and promoted shared decision making with conditional recommendations. For example, in patients with typical symptoms suggestive of a flare with only modestly elevated fecal calprotectin, and there is delay in performing endoscopic assessment due to logistical issues, patients and providers may be willing to initiate treatment, despite test performance suggesting FP rates of >5%.

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 multiple different cutoffs for different scenarios. Higher fecal calprotectin cutoffs may have modestly lower FP rates with modest improvement in confidence of decision making. In patients with typical symptoms suggestive of flare, an elevated CRP had very good performance, at least comparable with fecal tests. Although this may be convenient, it is important to note that stool testing to rule out *C difficile* and other enteric pathogens may still be required for all symptomatic patients.

Question 5: In patients with established UC, is interval biomarker-based monitoring superior to endoscopy-based monitoring to improve long-term outcomes?

Recommendation 8: In patients with UC, 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 UC 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 and those with elevated biomarkers would undergo endoscopy. In contrast, an endoscopy-based monitoring strategy involves routine endoscopic assessment to confirm achievement of endoscopic improvement (MES 0 or 1) or endoscopic remission (MES 0) target periodically. Supplementary Figure 8 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 UC 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. However, the performance of these biomarkers in a combination of symptoms may be more modest for detecting endoscopic remission (MES 0) and histologic remission, outcomes that have been associated with lower risk of clinical relapse compared with mild endoscopic activity (MES 1). 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 remission and histologic 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 UC in symptomatic remission. This was identified as a knowledge gap that warrants clinical trials.

Limitations of Current Evidence and Future Directions

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

Timing of measuring biomarkers. There were few studies that examined the accuracy and utility of serial measurements of serum or fecal biomarkers, particularly in settings where there was discordance between symptoms and biomarker values. As well, the optimal timing for this serial monitoring in either asymptomatic patients with UC or those with mild symptoms is unclear. In the post-induction setting under a treat-to-target paradigm, the optimal timing for measurement of biomarkers to inform treatment optimization has not been robustly established. In RCTs, biochemical response has been typically assessed 6–10 weeks after initiation of therapy. The STRIDE (Selecting Therapeutic Targets in Inflammatory Bowel Disease) consensus statement provides optimal time intervals for

assessment of clinical and endoscopic response to treatment; whether serum or fecal biomarkers follow a similar trajectory or whether there is benefit to earlier or more frequent assessment of biochemical response to guide therapy optimization remains a knowledge gap.[7]

Biomarker-based treat-to-target strategy in ulcerative colitis. In contrast to CD, for which treatment strategy trials, such as CALM, have demonstrated that incorporating biomarker assessment as part of the treat-to-target strategy is beneficial, there is a paucity of high-quality data confirming the value of a similar biomarker-based treat-to-target strategy in UC.[10] Indirect support for this is presented in the evidence synthesis where persistent biomarker elevation, despite being in symptomatic remission, is associated with a higher risk of relapse; however, direct evidence is lacking. Similarly, 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.

Prognostic significance of biomarkers. The guideline was focused on the performance of biomarkers for detecting moderate to severe endoscopic activity and did not examine prognostic significance of the magnitude and persistence of biomarker elevation. Most reviewed studies presented data on individual biomarkers and only provided performance around specific cutoffs, usually optimized for that study. 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 fecal calprotectin of 251 μg/g.[47] There were insufficient data to guide nuanced decision making in this context. Similarly, combination of biomarkers (elevated CRP and an elevated fecal calprotectin) in a given clinical setting may have different management implications than a single biomarker. 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.

Biomarker performance in diverse populations. Finally, the panel recognized the lack of robust data in specific clinical situations including mild UC, acute severe UC, and inflammatory disorders of the pouch, and in diverse patient populations where only a few studies examining the role of biomarkers to date exist.

What Do Other Guidelines Say?

There has been limited discussion on the role of noninvasive biomarkers in the management of UC in clinical guidelines. The American College of Gastroenterology Society guideline published in 2019 on the management of UC suggested fecal calprotectin as a surrogate for endoscopy when endoscopy is not feasible or available.[6] The European Crohn's and Colitis Organization and the European Society of Gastrointestinal and Abdominal Radiology guidelines on the diagnostic assessment of inflammatory bowel disease recognized that asymptomatic patients with elevated biomarkers of inflammation, mainly fecal calprotectin and CRP, may suggest imminent flare and recommended endoscopic or radiologic evaluation.[48] In patients with clinical response to medical therapy, the guidelines recommend evaluating for mucosal healing via either 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.

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 provide rapid guidance updates 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 https://doi.org/10.1053/j.gastro.2022.12.007

References

1. 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.
2. GBD 2017 Inflammatory Bowel Disease Collaborators. The global, regional, and national burden of inflammatory bowel disease in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol 2020;5:17–30.
3. Ungaro R, Mehandru S, Allen PB, et al. Ulcerative colitis. Lancet 2017;389:1756–1770.
4. 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.
5. Park KT, Ehrlich OG, Allen JI, et al. The cost of inflammatory bowel disease: an initiative from the Crohn's & Colitis Foundation. Inflamm Bowel Dis 2020;26:1–10.
6. Rubin DT, Ananthakrishnan AN, Siegel CA, et al. ACG clinical guideline: ulcerative colitis in adults. Am J Gastroenterol 2019;114:384–413.

7. 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.
8. Peyrin-Biroulet L, Ferrante M, Magro F, et al. Results from the 2nd Scientific Workshop of the ECCO. I: impact of mucosal healing on the course of inflammatory bowel disease. J Crohns Colitis 2011;5:477–483.
9. Turner D, Ruemmele FM, Orlanski-Meyer E, et al. Management of paediatric ulcerative colitis, part 1: ambulatory care-an evidence-based guideline from European Crohn's and Colitis Organization and European Society of Paediatric Gastroenterology, Hepatology and Nutrition. J Pediatr Gastroenterol Nutr 2018;67:257–291.
10. 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.
11. Ungaro RC, Colombel J-F, Yzet C, et al. Deep remission at 1 year prevents progression of early Crohn's disease. Gastroenterology 2020;159:139–147.
12. Limketkai BN, Singh S, Sandborn WJ, et al. US practice patterns and impact of monitoring for mucosal inflammation after biologic initiation in inflammatory bowel disease. Inflamm Bowel Dis 2019;25:1828–1837.
13. Yang JY, Lund JL, Pate V, et al. Utilization of colonoscopy following treatment initiation in U.S. commercially insured patients with inflammatory bowel disease, 2013-2019 [published online ahead of print August 22, 2022]. Inflamm Bowel Dis https://doi.org/10.1093/ibd/izac136.
14. 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.
15. Smalley W, Falck-Ytter C, Carrasco-Labra A, et al. AGA Clinical Practice Guidelines on the laboratory evaluation of functional diarrhea and diarrhea-predominant irritable bowel syndrome in adults (IBS-D). Gastroenterology 2019;157:851–854.
16. Carrasco-Labra A, Lytvyn L, Falck-Ytter Y, et al. AGA Technical Review on the evaluation of functional diarrhea and diarrhea-predominant irritable bowel syndrome in adults (IBS-D). Gastroenterology 2019;157:859–880.
17. 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.
18. Sands BE. Biomarkers of inflammation in inflammatory bowel disease. Gastroenterology 2015;149:1275–1285.e2.
19. 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.
20. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557–560.
21. 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.
22. 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.
23. Pathirana WPNGW, Paul Chubb SA, Gillett MJ, et al. Faecal calprotectin. Clin Biochem Rev 2018;39:77–90.
24. Axelrad JE, Joelson A, Green PHR, et al. Enteric infections are common in patients with flares of inflammatory bowel disease. Am J Gastroenterol 2018;113:1530–1539.
25. Limsrivilai J, Saleh ZM, Johnson LA, et al. Prevalence and effect of intestinal infections detected by a PCR-based stool test in patients with inflammatory bowel disease. Dig Dis Sci 2020;65:3287–3296.
26. Murthy SK, Feuerstein JD, Nguyen GC, et al. AGA Clinical Practice Update on endoscopic surveillance and management of colorectal dysplasia in inflammatory bowel diseases: expert review. Gastroenterology 2021;161:1043–1051.e4.
27. Adams A, Gupta V, Mohsen W, et al. Early management of acute severe UC in the biologics era: development and international validation of a prognostic clinical index to predict steroid response [published online ahead of print September 28, 2022]. Gut https://doi.org/10.1136/gutjnl-2022-327533.
28. Yoon H, Jangi S, Dulai PS, et al. Incremental benefit of achieving endoscopic and histologic remission in patients with ulcerative colitis: a systematic review and meta-analysis. Gastroenterology 2020;159:1262–1275.e7.
29. Gupta A, Yu A, Peyrin-Biroulet L, et al. Treat to target: the role of histologic healing in inflammatory bowel diseases: a systematic review and meta-analysis. Clin Gastroenterol Hepatol 2021;19:1800–1813.e4.
30. Kawashima K, Ishihara S, Yuki T, et al. Fecal calprotectin level correlated with both endoscopic severity and disease extent in ulcerative colitis. BMC Gastroenterol 2016;16:47.
31. Sonoyama H, Kawashima K, Ishihara S, et al. Capabilities of fecal calprotectin and blood biomarkers as surrogate endoscopic markers according to ulcerative colitis disease type. J Clin Biochem Nutr 2019;64:265–270.
32. Sakuraba A, Nemoto N, Hibi N, et al. Extent of disease affects the usefulness of fecal biomarkers in ulcerative colitis. BMC Gastroenterol 2021;21:197.
33. Yamamoto T, Shimoyama T, Matsumoto K. Consecutive monitoring of faecal calprotectin during mesalazine suppository therapy for active rectal inflammation in ulcerative colitis. Aliment Pharmacol Ther 2015;42:549–558.
34. D'Amico F, Peyrin-Biroulet L, Rubin DT, et al. International consensus on methodological issues in standardization of fecal calprotectin measurement in inflammatory bowel diseases. United Eur Gastroenterol J 2021;9:451–460.
35. Lasson A, Stotzer P-O, Isakssonc S, et al. The intra-individual variability of faecal calprotectin: a

prospective study in patients with active ulcerative colitis. J Crohns Colitis 2015;9:26–32.
36. Kristensen V, Malmstrom GH, Skar V, et al. Clinical importance of faecal calprotectin variability in inflammatory bowel disease: intra-individual variability and standardisation of sampling procedure. Scand J Gastroenterol 2016;51:548–555.
37. Calafat M, Cabre E, Manosa M, et al. High within-day variability of fecal calprotectin levels in patients with active ulcerative colitis: what is the best timing for stool sampling? Inflamm Bowel Dis 2015;21:1072–1076.
38. Du L, Foshaug R, Huang VW, et al. Within-stool and within-day sample variability of fecal calprotectin in patients with inflammatory bowel disease. J Clin Gastroenterol 2018;52:235–240.
39. Kluft C, de Maat MP. Genetics of C-reactive protein: new possibilities and complications. Arterioscler Thromb Vasc Biol 2003;23:1956–1959.
40. Brull DJ, Serrano N, Zito F, et al. Human CRP gene polymorphism influences CRP levels: implications for the prediction and pathogenesis of coronary heart disease. Arterioscler Thromb Vasc Biol 2003;23:2063–2069.
41. Jones J, Loftus EV Jr, Panaccione R, et al. Relationships between disease activity and serum and fecal biomarkers in patients with Crohn's disease. Clin Gastroenterol Hepatol 2008;6:1218–1224.
42. Moran CJ, Kaplan JL, Winter HS. Genetic variation affects C-reactive protein elevations in Crohn's disease. Inflamm Bowel Dis 2018;24:2048–2052.
43. Osterman MT, Aberra FN, Cross R, et al. Mesalamine dose escalation reduces fecal calprotectin in patients with quiescent ulcerative colitis. Clin Gastroenterol Hepatol 2014;12:1887–1893.e3.
44. Dulai PS, Sandborn WJ, Murphy J. Microsimulation model to determine the cost-effectiveness of treat-to-target strategies for ulcerative colitis. Clin Gastroenterol Hepatol 2021;19:1170.
45. Peura S, Fall T, Almqvist C, et al. Normal values for calprotectin in stool samples of infants from the population-based longitudinal born into life study. Scand J Clin Lab Invest 2018;78:120–124.
46. Velasco Rodriguez-Belvis M, Viada Bris JF, Plata Fernandez C, et al. Normal fecal calprotectin levels in healthy children are higher than in adults and decrease with age. Paediatr Child Health 2020;25:286–292.
47. 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.
48. 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. The authors would also like to thank Ms Kimberly Kan, patient representative, for providing a patient's perspective in developing these clinical guidelines.

Conflicts of interest
These authors disclose the following: 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). Ashwin N. Ananthakrishnan receives consulting fees from Menten AI and Iterative Scopes. Benjamin L. Cohen receives consulting fees from AbbVie, Celgene-Bristol Myers Squibb, Lilly, Pfizer, Sublimity Therapeutics, Takeda, TARGET RWE; CME Companies: Cornerstones, Vindico; Speaking: AbbVie; Educational Grant: Pfizer. Jeremy Adler received research grants from Janssen Research & Development, LLC. 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 Singh is supported by National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) grants K23DK117058 and R03DK129631. Dr Ananthakrishnan is supported by 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.