

Update from 2010 (standard operating procedure): protocol for the 2024 British Society of Gastroenterology Guidelines on colorectal surveillance in inflammatory bowel disease

Morris Gordon ,¹ Gaurav Bhaskar Nigam ,² Vassiliki Sinopoulou ,³ Ibrahim Al Bakir ,⁴ Adrian C Bateman ,⁵ Shahida Din ,^{6,7} Sunil Dolwani,⁸ Anjan Dhar ,^{9,10} O D Faiz,^{11,12} Bu Hayee ,¹³ Chris Healey,¹⁴ Christopher Andrew Lamb ,¹⁵ Simon Leedham,² Misha Kabir,¹⁶ Marietta Iacucci,^{17,18} Ailsa Hart ,¹⁹ A John Morris,²⁰ Marco Novelli,²¹ Tim Raine,²² Matt Rutter,²³ Neil A Shepherd,²⁴ Venkataraman Subramanian,²⁵ Nigel J Trudgill ,²⁶ Maggie Vance,¹² Ana Wilson,^{12,27} Lydia White,²⁸ Ruth Wakeman,²⁹ James E East²⁸

To cite: Gordon M, Nigam GB, Sinopoulou V, *et al.* Update from 2010 (standard operating procedure): protocol for the 2024 British Society of Gastroenterology Guidelines on colorectal surveillance in inflammatory bowel disease. *BMJ Open Gastroenterol* 2024;**11**:e001541. doi:10.1136/bmjgast-2024-001541

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/bmjgast-2024-001541>).

Received 29 July 2024
Accepted 18 September 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Dr Morris Gordon;
MGordon@uclan.ac.uk

ABSTRACT

Introduction The evolving landscape of inflammatory bowel disease (IBD) necessitates refining colonoscopic surveillance guidelines. This study outlines methodology adopted by the British Society of Gastroenterology (BSG) Guideline Development Group (GDG) for updating IBD colorectal surveillance guidelines.

Methods and analysis The ‘Grading of Recommendations, Assessment, Development and Evaluation’ (GRADE) approach, as outlined in the GRADE handbook, was employed. Thematic questions were formulated using either the ‘patient, intervention, comparison and outcome’ format or the ‘current state of knowledge, area of interest, potential impact and suggestions from experts in the field’ format. The evidence review process included systematic reviews assessed using appropriate appraisal tools. An extensive list of potential outcomes was compiled from literature and expert consultations and then ranked by GDG members. The top outcomes were identified for evidence synthesis in three key areas: utility of surveillance in IBD, quality of bowel preparation and use of advanced imaging techniques in colonoscopy for IBD. Risk thresholding exercises determined specific risk levels for different surveillance strategies and intervals. This approach enabled the GDG to establish precise thresholds for interventions based on relative and absolute risk assessments, directly informing the stratification of surveillance recommendations. Significance of effect sizes (small, moderate, large) will guide the final GRADE assessment of the evidence.

Ethics and dissemination Ethics approval is not applicable. By integrating clinical expertise, patient experiences and innovative methodologies like risk thresholding, we aim to deliver actionable recommendations for IBD colorectal surveillance. This

WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ Patients with colonic inflammatory bowel disease (IBD) have an elevated risk of dysplasia and colorectal cancer (CRC), though this has decreased over time.
- ⇒ Limited evidence on surveillance strategies, biomarker use and dysplasia management, along with advancements in endoscopy and personalised care, highlights the need for updated guidelines and further research.

WHAT THIS STUDY ADDS

- ⇒ The Grading of Recommendations, Assessment, Development and Evaluation methodology guides the development of IBD surveillance guidelines, focusing on outcome selection and risk thresholds. A diverse expert group ensures a comprehensive and evidence-based approach.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ Updated guidelines will provide clear recommendations for dysplasia and CRC surveillance in IBD, potentially influencing policy and standardising resource-efficient strategies.
- ⇒ The focus on outcome selection and risk thresholds may inform future research and study design, advancing IBD surveillance and treatment.

protocol, complementing the main guidelines, offers GDGs, clinical trialists and practitioners a framework to inform future research and enhance patient care and outcomes.

INTRODUCTION

Inflammatory bowel disease (IBD), comprising ulcerative colitis (UC), Crohn's disease (CD) and IBD-unclassified (IBD-U), is a chronic and debilitating condition affecting a significant portion of the population worldwide, and its prevalence continues to rise. The worldwide prevalence of IBD increased from an estimated 3.32 million cases in 1990 to 4.90 million cases in 2019, marking a 47.45% increase over this period.¹ One of the major long-term complications of IBD is the increased risk of colorectal cancer (CRC), particularly in patients with long-standing and extensive colonic disease. Dysplasia, a precancerous condition, plays a crucial role in the progression to CRC in most IBD patients. Early detection of CRC and management of dysplasia are essential for improving patient outcomes and reducing the burden of CRC in this vulnerable population.²

The British Society of Gastroenterology (BSG) has been at the forefront of providing evidence-based clinical guidelines for the management of various gastrointestinal disorders, including IBD. As medical knowledge and technologies continue to advance, there is an increasing need to revisit and update existing guidelines to ensure that healthcare professionals have access to the most current and accurate recommendations.

In 2010, the BSG published formal guidance on this subject.³ In 2019, the BSG IBD guidelines provided further concise guidance on this topic (pages 70–71 and Box 11), consolidating the BSG 2010 guidelines with the North American SCENIC 2015 (Surveillance for Colorectal Endoscopic Neoplasia Detection and Management in Inflammatory Bowel Disease Patients: International Consensus Recommendations) guidelines.^{3–5} However, it did not initiate new systematic reviews or voting. This guideline, therefore, will represent a formal update of the 2010 guideline while also incorporating updates from the 2019 position. It aims to be co-published alongside the primary BSG 2024 IBD guidelines.⁶

This document serves as a protocol, outlining the technical review methods and a broader set of operating procedures that have been prospectively agreed to develop the updated guidelines. It comprehensively covers the multidisciplinary approaches employed in contemporary clinical practice. The final guideline will feature the official recommendations of the BSG Guideline Development Group (GDG) on aspects of colonoscopic surveillance for IBD patients. It is designed to support patients and professionals across various treatment settings and, as such, will be presented systematically and transparently in accordance with the best international methods guidance. The prospective publication of this document aligns with the high-quality standards being upheld throughout the process.⁷

METHODS AND ANALYSIS

The production of this guideline adheres to the procedures outlined in the 'Grading of Recommendations,

Assessment, Development and Evaluation' (GRADE) approach, as detailed in the GRADE handbook, and supported by the WHO handbook for guideline development.^{8,9} These resources provide comprehensive methods for every aspect of the guideline development process. The team will employ the GIN-McMaster guideline development checklist, an 18-point process map, to support each of these steps within a GRADE-compliant guideline development process.^{10,11}

Organisation, planning and training

In July 2023, members of the BSG IBD surveillance GDG held their inaugural meeting. The framework for its development was established, with technical review responsibilities assigned to a team at the Biomedical Evidence Synthesis and Translation research unit at the University of Central Lancashire, Preston (MG, VS), in collaboration with the team at the Translational Gastroenterology Unit, Nuffield Department of Medicine, Oxford (JE, GN). This team will oversee tasks such as searches, table creation and result synthesis. The joint guideline chairs will consist of a content and field expert (JE) and will be accompanied by a lead GRADE methodologist as co-chair (MG), who will have a non-voting role, as per GRADE procedures.¹² Administrative support will be available from both host Higher Education Institutions of the co-chairs, as well as access to a Cochrane expert information specialist arranged through these institutions.

Guideline development group (GDG) membership

The GDG is composed of members from BSG, including general and specialist endoscopists, gastroenterologists specialising in IBD, a nurse endoscopist, an IBD nurse specialist, specialist gastrointestinal pathologists, an IBD surgeon, a trainee representative, and patient and IBD stakeholder representatives, from across the UK. All members of the GDG will have voting rights within the GDG, while the methodological chair and methodological core team will remain a non-voting members.

GDG members were selected following criteria set by the BSG's Clinical Services and Standards Committee (CSSC), ensuring representation of a wide range of expertise, experience, views and skillsets with appropriate consideration for diversity, equity and inclusion. The planned group was submitted to CSSC for confirmation of meeting BSG criteria prior to the first working group meeting. The guideline development process adheres to the Appraisal of Guidelines for Research and Evaluation II principles of transparency, rigour and inclusion.¹³ Furthermore, BSG guidelines are accredited by NICE (National Institute for Health and Care Excellence), reflecting adherence to these high standards. The selection process was based on the CSSC criteria to ensure decisions were made through a structured approach, rather than solely by the chair.

All members of the team will be invited to be co-authors of the full guideline. They will be committed to maintaining the confidentiality of open discussions and

debates within the guideline process, as well as the confidentiality of the guideline's content before publication. Conflict of interest declarations were required from all members and will be reviewed throughout the process to maintain transparency and prevent bias.

Guideline development group (GDG priority setting and identifying target audience)

A key consideration involves prioritising stakeholders' perspectives concerning specific clinical or patient factors. To address this, the GDG members have been selected with significant national and international expertise in developing guidelines within the topic area. The inclusion of patients on the team is essential to ensure a wide range of viewpoints are represented. Patient perspectives and preferences play a central role in guideline development. The updated guidelines will prioritise patient-centred care, considering the individual needs, values and expectations of IBD patients in dysplasia surveillance decisions.

The team convened in July 2023, using previous guidelines as a foundation to identify broad thematic questions. The starting list of questions covered under six broad themes are detailed in online supplemental file 1. A final consensus list of thematic questions will be agreed on before the technical review phase. This process will refine the questions into the PICO format, which will guide the review of relevant trials and observational data as described below. For certain questions that do not align with this format, we will use the Current state of knowledge, Area of interest, Potential impact and suggestion from experts in the field (CAPS) formulation to transition from justification-based PICO questions to more descriptive or clarification questions.¹⁴

It is important to note that the PICOs and CAPS presented in the online supplemental file 1 are not final. They represent the initial framework developed primarily from the statements and questions in the BSG 2010 guidelines and the SCENIC 2015 statement that will be further developed and refined during the guideline development process, allowing for the inclusion of new evidence and additional areas of interest.^{3 4} The GDG remains flexible throughout the process to ensure comprehensive coverage of all relevant issues.

Stages of process

The following fundamental procedures will govern the main stages of the guideline development:

- ▶ The prospective publishing of a guideline protocol and technical summary protocol in an open access journal (this manuscript).
- ▶ Prospective agreement of thresholds for risk and methods for stratifying these risk categories, prior to production of technical review output.^{15 16}
- ▶ The completion of a detailed, methodologically rigorous technical review which will include GRADE summary of findings for all outcomes and preparation of evidence to decision (ETD) frameworks for

PICO questions to support the GDG decision-making, as well as detailed narrative evidence summaries for other questions.¹⁷

- ▶ A face-to-face summit of the GDG to discuss the evidence within the ETD and summaries. This will be followed by anonymous voting and further discussion to reach a consensus on items with disagreements.
- ▶ The publication of a concise main guideline that summarises key recommendations, the certainty of underpinning evidence and the strength of the recommendations, all within the main published journal output.
- ▶ An accompanying patient and public focused decision-making aid version of the guideline to support practical and autonomous coproduction of treatment plans.

This series of outputs offers systematic, high quality and high utility output for all our audiences.

Patient, intervention, comparison and outcome (PICO) question generation

The generation of questions will be guided by the GRADE guidelines.¹⁸

Key areas of focus for refinement of all PICO questions will be considered by the GDG. These core elements of refinement around PICO questions and their specific application will be presented in draft form to the GDG and all feedback considered, with the final list below:

- ▶ Multiple intervention arms will be considered. To allow consideration of non-placebo comparators and standard therapies, network meta-analysis will be deployed in key targeted areas, as decided by the GDG and when sufficient volume of similar studies exist. Subgroup analyses will be performed for outcome measures in the case of different comparator groups, given that heterogeneity and sufficient volume of studies exists. This is expected to be limited within the scope of the guideline.
- ▶ Any context of surveillance with patients suffering from either ulcerative colitis, colonic Crohn's disease or IBD-unclassified will be considered. Patients with microscopic colitis and isolated small bowel Crohn's disease will not be included.

Threshold and risk stratification prospective agreement

In addressing the thematic focus on factors influencing surveillance decisions in IBD, we synthesised evidence to support risk stratification and to inform targeted surveillance recommendations. The GDG adopted a structured approach, combining published guidelines with the collective expertise of its members to establish consensus and set appropriate thresholds.¹⁹

Risk stratification methodology

The GDG employed online questionnaires (JotForm) to determine the risk thresholds for CRC in IBD patients compared with the general population. These thresholds define the risk levels at which surveillance becomes

Table 1 Risk thresholding for surveillance frequency categorisation

	Small mean (SD)	Moderate mean (SD)	Large mean (SD)
Risk threshold points at which the transition occurs from trivial to small risk, small to moderate risk and moderate to large risk for patients who have entered the surveillance pathway with colonoscopy			
Relative risk	1.7 (0.5)	3 (1.3)	5.2 (3.1)
Absolute risk	3.6% (2.9)	6.9% (4.4)	14.4 (9.1)

necessary.²⁰ Based on relative and absolute risks, the GDG suggested frequency intervals for surveillance, categorising patients into low, medium and high-risk groups for developing IBD-associated advanced colorectal neoplasia.

Outcome selection and ranking

An extensive list of potential outcomes was compiled from the literature and expert consultations. GDG members ranked their top seven outcomes in order of importance. These rankings were used to calculate a cumulative score to identify the most critical outcomes for evidence synthesis in the following thematic areas:

1. Utility of surveillance in IBD.

2. Quality of bowel preparation in IBD colonoscopy.
3. Use of advanced imaging techniques in colonoscopy for IBD.

Risk thresholding and effect size determination

Additional risk thresholding exercises were conducted focusing on colonoscopic modalities and the quality of bowel preparation. These exercises were designed to determine the significance of effect sizes (small, moderate, large) as perceived by GDG members, based on randomised controlled trial (RCT) data. This process is essential for the final GRADE assessment of the evidence.

The GDG established an average relative risk (RR) of 1.5 (SD 0.4) for CRC in IBD compared with the general population as a criterion for considering colonoscopic surveillance. The outcomes selected, along with results from the risk thresholding, are detailed in [tables 1 and 2](#) and [box 1](#). Additional information, including response rates and median and IQR values for each question, is provided in the online supplemental file 2.

The risk thresholding exercise was conducted anonymously among GDG members; however, members were not required to vote and were recommended to vote only on areas where they felt comfortable that they had sufficient expertise to contribute. The results, presented in the online supplemental file 2, revealed variability

Table 2 Outcomes and risk thresholds for quality of bowel preparation and use of colonoscopic modalities (advanced imaging techniques) in IBD

	Small mean (SD)	Moderate mean (SD)	Large mean (SD)
Quality of bowel preparation			
Preparation quality (using validated scores)	6.7% (3)	12.7% (7.7)	23.5% (14.7)
Adenoma/polyp detection rates	3.9% (2.8)	7.2% (4.4)	12.3% (7.5)
Patient tolerability to take/complete the bowel prep	5.6% (2.9)	11.2% (7.4)	18.8% (11.6)
Patients with serious adverse events only	2.4% (1.3)	4% (2.8)	6.4% (5.2)
Caecal intubation rates	3.5% (1.5)	6.9% (3.6)	10.8% (5.4)
Patient acceptability/willingness to repeat	4.9% (2.8)	10.7% (7.2)	17.4% (11.4)
Patient withdrawals due to adverse events	3.6% (2.5)	5.1% (3.2)	9.3% (8)
Colonoscopic modalities (advanced imaging techniques) in IBD			
Detection of dysplastic lesions (as per Vienna classification: indefinite for dysplasia, low-grade dysplasia, high-grade dysplasia or invasive neoplasia at histological examination)	3.3% (2.4)	5.8% (3)	11.2% (7.1)
Yield of any dysplasia from targeted biopsies (per patient)	3.4% (2.9)	6.7% (5)	10.9% (7.5)
Yield of any dysplasia from random biopsies (per patient)	3.5% (4.8)	6.2% (7.2)	10% (10.2)
Patients with serious adverse events	2.6% (2.5)	5.1% (4.7)	8.4% (7.1)
Detection of any lesions in patients (neoplastic lesions detected, that is, dysplastic+serrated and/or non-neoplastic-endoscopic findings with no evidence of dysplasia or invasive neoplasia at histology)	4.1% (2.2)	7.9% (4.4)	15.1% (12.4)
Patient acceptability/willingness to repeat	3.7% (2.4)	6.1% (4.9)	9.6% (7.5)
Patient withdrawals due to adverse events	3.1% (2.5)	5.5% (4.8)	8.6% (7.4)

IBD, inflammatory bowel disease.

Box 1 Outcomes selected for utility of colonoscopic surveillance in inflammatory bowel disease

Colorectal cancer (CRC) detection
 Death/survival related to CRC
 Tumour stage (early/late) detection of CRC
 Patients with serious adverse events only
 Rates of missed CRCs
 Rates of colectomy/surgical resections
 Adherence to surveillance by healthcare professional

in responses, as indicated by wide SD in some areas—a reflection of the diverse clinical practices represented. While this is a limitation, it also marks progress in standardising risk assessment. Achieving tighter confidence intervals may require a larger group of international experts and stakeholders, a goal for potential future collaborations beyond the current guideline update.

This approach ensures that the synthesis of evidence does not lead to biased decision-making based on the strength or magnitude of the results. Instead, the evidence will be interpreted within the context of the *a priori* framework.

This approach is innovative for such a guideline but is built on the method used to establish thresholds within GRADE guidelines.¹⁵

Evidence selection

Types of studies

We will include all published, unpublished and ongoing RCTs that compare interventions with other active interventions, standard therapy, placebo or no therapy. Studies that do not report any of the outcome measures of interest will be excluded. In the case of diagnostic test accuracy questions or epidemiological questions we will also consider observational studies.

Types of participants

Adult patients >18 years of age with a diagnosis of either ulcerative colitis, colonic Crohn's disease, or IBD-unclassified defined by conventional clinical, endoscopic and histologic criteria who would be considered eligible for surveillance, based solely on the duration and extent of disease.

Types of outcome measures

Both dichotomous and continuous outcomes will be considered, as per the appropriate questions.

Search methods for identification of studies

Electronic searches

We will use a search strategy designed and checked by an information specialist with Cochrane expertise (online supplemental file 3)

We will search: the Cochrane Central Register of Controlled Trials (CENTRAL) (via Ovid EBMR) (inception to present); MEDLINE (via Ovid) (1946 to present); Embase (via Ovid) (1974 to present); PsycINFO (via

Ovid) (1987 to present); AMED (via Ovid) (Allied and Complementary Medicine) (1985 to present); and CINAHL (via EBSCO) (Cumulative Index to Nursing and Allied Health Literature) (1984 to present).

We will place no restrictions on language of publication.

Searches will be produced for each of the specific PICO and non-PICO/CAPS-based questions to appropriately include studies.

A three-phase approach will be employed for searching for studies.

1. Systematic reviews will be included. Potential will be assessed using the Assessment of Multiple Systematic Reviews (AMSTAR 2) tool.²¹ When multiple reviews are found on the same topic, the highest rated review will be included. Assessors in pairs will determine if the AMSTAR rated reviews are of sufficient quality to be included, with consensus on ratings reached by a third assessor. If these studies are not up to date (completed within the last 18 months) or if any additional studies are identified in the broader search or from rejected systematic reviews, they will be incorporated, and the meta-analysis will be rerun to update the results. In cases where risk of bias or GRADE ratings are not included as needed, they will be addressed using the approach mentioned below. Cochrane systematic reviews will be given priority for inclusion, subject to the same conditions of updating analyses to encompass all relevant studies.
2. If appropriate for the question, RCTs that assess the interventions of interest will be included for consideration. Phase 1 studies will not be included. Only randomised trials will be included; quasi-randomised or non-randomised studies will not be considered. These studies will be extracted and analysed in accordance with the methods outlined below and, when applicable, combined with the systematic reviews mentioned earlier. Quality assessment of all RCTs will be conducted using the risk of bias tool.
3. If appropriate for the question, other observational study designs will be considered. These will be assessed using the Risk Of Bias In Non-randomised Studies - of Interventions tool for risk of bias or diagnostic test accuracy using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool.^{22 23}

Data collection and analysis

We will carry out data collection and analysis according to the methods recommended in the Cochrane Handbook for Systematic Reviews of Interventions.²⁴

Selection of studies

Two or more authors will independently review the titles and abstracts identified through the literature search, excluding studies that, based on their titles and abstracts, are not relevant. All reviews will be conducted in duplicate independently, and any disagreements will be resolved through consensus with a third author. Full reports of studies deemed potentially eligible will be obtained.

These reviewers will independently assess the full texts for inclusion in the review, and any disagreements will again be resolved through discussion with a third author. We will document the studies excluded at this or subsequent stages, along with the primary reason for their exclusion.

In cases where there are multiple publications for a given study, we will compile the reports of the same study.

Data extraction and management

Authors will independently perform data extraction using piloted data extraction forms. We will extract the following data from the included studies:

- ▶ Study setting: country and number of trial centres.
- ▶ Methods: study design, total study duration and date.
- ▶ Participant characteristics: age, sociodemographics, ethnicity, diagnostic criteria and total number of participants.
- ▶ Eligibility criteria: inclusion and exclusion criteria.
- ▶ Intervention and comparator.
- ▶ Outcomes: outcome definition, unit of measurement and time of collection.
- ▶ Results: number of participants allocated to each group, missing participants and sample size.
- ▶ Funding source.

Assessment of risk of bias in included studies

Risk of bias in the included RCT studies will be independently assessed by two or more authors, based on the criteria outlined in the Cochrane Handbook for Systematic Reviews of Interventions.²⁴ Where feasible, we will contact lead authors of included studies to determine the true risk of bias.

We will assess the following 'risk of bias' domains:

- ▶ Sequence generation (selection bias).
- ▶ Allocation concealment (selection bias).
- ▶ Blinding of participants and personnel (performance bias).
- ▶ Blinding of outcome assessment (detection bias).
- ▶ Incomplete outcome data (attrition bias).
- ▶ Selective reporting (reporting bias).
- ▶ Other biases such as imbalance in participants' baseline characteristics.

The studies will be judged to be at low, high or unclear risk of bias for each domain assessed, based on the guidance in the Cochrane Handbook for Systematic Reviews of Interventions.²⁴

After data extraction, the review authors will compare the extracted data, discussing and resolving any discrepancies before transfer of data into the 'Characteristics of included studies tables.

For diagnostic accuracy studies, the QUADAS tool will be used with the following items considered²⁵:

1. Representative spectrum.
2. Acceptable reference standard.
3. Acceptable delay between tests.

4. Partial verification avoided.
5. Differential verification avoided.
6. Incorporation avoided.
7. Index test results blinded.
8. Reference standard results blinded.
9. Relevant clinical information.
10. Uninterpretable results reported.
11. Withdrawals explained.

Measures of treatment effect

We will express treatment effect as RRs with corresponding 95% CIs for dichotomous outcomes, and mean difference (MD) with 95% CI for continuous outcomes. Where endpoint and change score were both reported, we will use endpoint scores for data analysis. However, if the studies assessed the same continuous outcome in different ways, we would estimate the treatment effect using the standardised mean difference (SMD).²⁴

Unit of analysis issues

The unit of analysis will be the participants. In studies comparing more than two intervention groups, we intend to perform multiple pairwise comparisons between all possible pairs of intervention groups. To prevent double counting, we will evenly distribute shared intervention groups among these comparisons. For dichotomous outcomes, we plan to divide both the number of events and the total number of participants. For continuous outcomes, we will only divide the total number of participants, keeping the means and SDs unchanged.

Cross-over studies will be included in quantitative analysis only if data are reported separately for before and after the cross-over, using pre-cross-over data exclusively. We do not anticipate encountering any cluster RCTs; however, if such trials are identified, we will only use their data if the authors have employed appropriate statistical methods to account for the clustering effect. In a sensitivity analysis, we will also exclude cluster RCTs to evaluate their impact on the results.

Dealing with missing data

In instances of missing data or studies that have not reported data in sufficient detail, we will proactively reach out to study authors. We will make efforts to estimate missing SDs using appropriate statistical tools and calculators available within Review Manager 5 if the studies have reported standard errors (Review Manager 2020). Studies that do not provide measures of variance will be considered at high risk of reporting bias.

Assessment of heterogeneity

We will assess the included studies to evaluate their homogeneity in terms of participants, intervention, comparator and outcome. To assess statistical heterogeneity, we will use a χ^2 test with a significance level set at $p < 0.1$ to indicate the presence of heterogeneity. Inconsistency will be quantified and expressed through the I^2 statistic. We will interpret the thresholds as follows:²⁴

- ▶ 0% to 40%: might not be important.

- ▶ 30% to 60%: may represent moderate heterogeneity.
- ▶ 50% to 90%; may represent substantial heterogeneity.
- ▶ 75% to 100%: considerable heterogeneity.

Assessment of reporting biases

Most reporting biases can be mitigated using an inclusive search strategy. We intend to explore the possibility of publication bias by employing a funnel plot when we have 10 or more studies available for analysis. The extent of publication bias will be assessed through visual examination of funnel plot asymmetry.

Additionally, we will test funnel plot asymmetry by conducting a linear regression of the intervention effect estimate against its SE, applying weighting based on the inverse of the variance of the intervention effect estimate.²⁶

Data synthesis

To summarise the characteristics of the included studies, we will initially conduct a narrative synthesis encompassing all of them. Subsequently, we will perform a meta-analysis if two or more studies have assessed similar populations, interventions and outcomes. We plan to analyse studies involving children, adults and different sub-intervention types separately.

We will use Review Manager 5 (Review Manager 2020) for our data synthesis. The random-effects model will be used to combine study data. Effect estimates from studies reporting data in a similar manner will be pooled in the meta-analysis. For dichotomous outcomes, we will pool RRs, whereas for continuous outcomes, we will pool MDs or SMDs. These results will be presented alongside 95% CIs.

In cases where conducting a meta-analysis is not feasible, typically due to variations in data reporting, we will provide a narrative summary of the included studies.

Subgroup analysis and investigation of heterogeneity

If we identify heterogeneity, we will investigate possible causes and address them using the methods described in the Cochrane Handbook for Systematic Reviews of Interventions.²⁴ We plan to undertake subgroup analyses of potential effect modifiers if sufficient data were available.

Sensitivity analysis

We plan to conduct a sensitivity analysis focused on the primary outcome of treatment success to assess the robustness of our review findings concerning decisions made during the review process. This analysis will involve excluding studies with a high or unclear risk of bias from our analyses. In instances where data analyses include studies with both reported and estimated SDs, we will exclude studies with estimated SDs to examine the impact on our review's findings. Furthermore, we will explore whether the choice of model (fixed-effect vs random-effects) influences the results. In cases of unexplained heterogeneity, we will perform a targeted investigation of key factors within any outlier studies to better understand and potentially define the source of this heterogeneity.

Summary of findings and assessment of the certainty of the evidence

We will present our primary outcomes results in 'Summary of findings' tables for all forms of studies. For PICO questions, we will export to GRADEpro GDT software for quality assessment (GRADEpro GDT).²⁷ Based on risk of bias, inconsistency, imprecision, indirectness and publication bias, we will grade the quality of the evidence for each outcome as high, moderate, low or very low. This will use the targeted and outcome specific thresholds to support imprecision judgements. These ratings have been defined as follows:

- ▶ High: further research is very unlikely to change our confidence in the estimate of effect.
- ▶ Moderate: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
- ▶ Low: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
- ▶ Very low: any estimate of effect is very uncertain.

We will justify all decisions to downgrade the quality of studies using footnotes and make comments to aid the reader's understanding of the review where necessary.

Thresholds of treatment effect

When comparing two interventions or approaches, all RRs will be supplemented with absolute risk difference and appropriate confidence intervals of absolute effects. These will be categorised according to the thresholds that have been defined by the GDG to aid interpretation of the clinical significance of the finding.

Development of recommendations

The complete technical summary will be provided to voting members after conducting an updated search to incorporate any new studies and integrate them into the existing evidence. The data and GRADE summary of findings tables will be incorporated into ETD frameworks,²⁸ facilitating the consideration of key factors to inform decision-making.

In cases where evidence is limited, we will provide recommendations using the GRADE 'expert evidence approach'.²⁹ For questions that do not follow the PICO format but are descriptive in nature, we will present a narrative summary to support best practice statements or similar formulations.

A face-to-face meeting will be convened to thoroughly discuss, explore and critically evaluate the components of the completed technical review and the ETD frameworks. When clear agreement is reached, recommendations will be prepared, followed by anonymous voting to confirm consensus. In instances of disagreement, the ETD framework will guide the voting process and help identify the underlying reasons for such disagreement. The team will then meet to discuss these findings and endeavour to formulate any relevant consensus recommendations.

Any unresolved disagreements will also be included in the guideline discussion.

Voting will be based on a clear GRADE statement with accompanying justification and implementation statements, along with magnitude and certainty data. The votes will be dichotomous (Yes or No) and must reach 75% agreement to approve the item. If an agreement is not reached, further discussion will be conducted, amendments made, and both the original and amended statements will be voted on sequentially. If neither attains 75%, the discussion will be temporarily halted and resumed later in the day to allow time for reflection. The team will gather and refocus the evidence, followed by another round of discussion.

Good practice and narrative items will be discussed, refined and a broad consensus reached, but formal voting will not be conducted.

The non-voting team will refine these recommendations into a final list, ensuring that the strength of the recommendations aligns with the presented evidence and the views of the GDG, in accordance with GRADE recommendation guidance. The final proposals will be agreed on by consensus, with the strength of agreement, certainty of evidence and strength of recommendations all clearly presented. The synthesised recommendations will be prepared in a guideline that adheres to BSG and journal publication standards. The ETD frameworks will be made available as supplementary material, and the technical evidence will be published in full as accompanying outputs to support the primary guidance.

Areas of future research

During the development of this guideline, we will identify key areas in need of further research that will facilitate future priority setting partnerships.

Ethics and dissemination

Ethics approval is not applicable. By integrating clinical expertise, patient experiences and innovative methodologies like risk thresholding, we aim to deliver actionable recommendations for IBD colorectal surveillance. This protocol, complementing the main guidelines, offers GDGs, clinical trialists and practitioners a framework to inform future research and enhance patient care and outcomes.

Author affiliations

- ¹BEST unit, University of Central Lancashire, Preston, UK
- ²University of Oxford Translational Gastroenterology Unit, Oxford, UK
- ³University of Central Lancashire, Preston, UK
- ⁴Gastroenterology Department, Chelsea and Westminster Hospital, London, UK
- ⁵Cellular Pathology, University Hospital Southampton NHS Foundation Trust, Southampton, UK
- ⁶Gastroenterology, NHS Lothian, Edinburgh, UK
- ⁷NHS Research Scotland, Clydebank, UK
- ⁸Gastroenterology, Cardiff and Vale NHS Trust, Cardiff, UK
- ⁹Gastroenterology, County Durham and Darlington NHS Foundation Trust, Darlington, UK
- ¹⁰School of Health and Life Sciences, Teesside University, Middlesbrough, UK
- ¹¹Department of Surgery and Cancer or Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK

- ¹²St Mark's Hospital and Academic Institute, London, UK
- ¹³King's College Hospital NHS Foundation Trust, London, UK
- ¹⁴Gastroenterology, Airedale NHS Foundation Trust, Keighley, UK
- ¹⁵Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK
- ¹⁶University College London, London, UK
- ¹⁷University College Cork, Cork, Cork, Ireland
- ¹⁸University of Birmingham, Birmingham, UK
- ¹⁹St Mark's Hospital, London, UK
- ²⁰Gastroenterology, Glasgow Royal Infirmary, Glasgow, UK
- ²¹Department of Histopathology, University College London, London, UK
- ²²Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
- ²³University Hospital of North Tees, Stockton-on-Tees, UK
- ²⁴Gloucestershire Hospitals NHS Foundation Trust, Cheltenham, UK
- ²⁵Gastroenterology, Leeds Teaching Hospitals NHS Trust, Leeds, UK
- ²⁶Department of Gastroenterology, Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK
- ²⁷Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK
- ²⁸Nuffield Department of Medicine, University of Oxford Translational Gastroenterology Unit, Oxford, UK
- ²⁹Crohn's & Colitis UK, Hatfield, UK

X Gaurav Bhaskar Nigam @drgauravnigam, Shahida Din @ShahidaDin1, Anjan Dhar @anjan_dhar6 and Christopher Andrew Lamb @DrChrisLamb

Acknowledgements We are grateful for the contribution of our patient representatives Caroline Brocklehurst and Vicky Buckley. GBN is supported by a National Institute for Health and Care Research (NIHR) Doctoral Research Fellowship (Grant number 302607). JEE is supported by the NIHR Oxford Biomedical Research Centre.

Contributors MG and GBN co-wrote the manuscript and are joint first authors. MG and VS provided input on methodological aspects. GBN wrote the questions for risk thresholding, supervised by VS, MG and JEE. Main supervision was provided by MG and JEE. All authors (MG, GBN, VS, IAB, AB, SDin, SDoI, AD, ODF, BH, CH, CAL, SL, MK, MI, AH, AJM, MN, TR, MR, NAS, VS, NJT, MV, AW, LW, RW, JEE) reviewed and approved the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests MG is an editor for Cochrane, supporting clinical truth and unbiased synthesis since 2017 (ongoing). GBN, VS, IAB, AB, SDO, OF, BH, JM, MN, NS, VSU and JEE have no competing interests to declare related to the manuscript. SD is a committee member of the BSG IBD Section (2019–2025) and National Lead for IBD Colorectal Cancer Surveillance for the Scottish Government's National Centre for Sustainable Delivery (2020–2025). AD has received speaker fees and honoraria from AbbVie (2023–present), with his department receiving advisory fees, speaker fees and consultancy support from Tillotts UK (since 2022) and speaker fees, honoraria and advisory support from Takeda (since 2022). CH is employed as a JAG assessor, with JAG using standards to guide the assessment of endoscopy units for accreditation. CL has received research support and/or fees for non-promotional education from Janssen, Takeda, AbbVie, AstraZeneca, Eli Lilly, Orion, Pfizer, Roche, Sanofi Aventis, Ferring, UCB, Nordic Pharma, Dr Falk and Biogen (2019–present), as well as research support from Genentech Inc. (2012–present); he is also the Secretary of the IBD Section for the British Society of Gastroenterology (2021–present). AH has received lecture and advisory board fees from Janssen, Takeda, Celltrion, Galapagos, AbbVie and Lilly (2017–present), and her department has received sponsorships for meetings from Janssen, Takeda, Celltrion, Galapagos and AbbVie (2020–present); she is also a Governing Board member of ECCO (since 2018) and a Scientific Board member of Crohn's and Colitis UK (since 2011). TR has undertaken paid consultancy work for Abbvie, Arena, Aslan, AstraZeneca, Boehringer-Ingelheim, BMS, Eli Lilly, Ferring, Galapagos, Gilead, GSK, Heptares, LabGenius, Janssen, Pfizer, Roche, Takeda, UCB and XAP Therapeutics (since 2020), and he is a director of Endoread Ltd. (since 2022); he serves on the Scientific Committee of ECCO (since 2022) and United European Gastroenterology (since 2021). MR is a Regional Endoscopy Advisor for NHS England (since 2021) and Chair of JAG (since 2023). NT serves on the advisory board for Dr Falk UK concerning EoE and Jorveza (current), and his department receives research support, speaker fees and conference support from Dr Falk UK (research support is current). LW has served on advisory boards for Pfizer (May 2023), BMS (June 2023), Celltrion (June 2023) and Dr Falk (June 2023); she has

received a speaker's honorarium from Takeda (February 2023) and conference attendance support from Dr Falk (June 2023); she also receives editorial royalties for the Inflammatory Bowel Disease Nursing Manual (2019). RW is an employee of Crohn's & Colitis UK (since October 2019) and Crohn's & Colitis UK receives funding from various sources, as detailed in their annual accounts (published online).

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Morris Gordon <http://orcid.org/0000-0002-1216-5158>

Gaurav Bhaskar Nigam <http://orcid.org/0000-0003-4699-2263>

Vassiliki Sinopoulou <http://orcid.org/0000-0002-2831-9406>

Ibrahim Al Bakir <http://orcid.org/0000-0001-8775-3459>

Adrian C Bateman <http://orcid.org/0000-0003-2222-4104>

Shahida Din <http://orcid.org/0000-0003-2855-3400>

Anjan Dhar <http://orcid.org/0000-0001-8964-2031>

Bu Hayee <http://orcid.org/0000-0003-1670-8815>

Christopher Andrew Lamb <http://orcid.org/0000-0002-7271-4956>

Ailsa Hart <http://orcid.org/0000-0002-7141-6076>

Nigel J Trudgill <http://orcid.org/0000-0002-8040-8158>

REFERENCES

- Wang R, Li Z, Liu S, *et al*. Global, regional and national burden of inflammatory bowel disease in 204 countries and territories from 1990 to 2019: a systematic analysis based on the Global Burden of Disease Study 2019. *BMJ Open* 2023;13:e065186.
- Zisman TL, Rubin DT. Colorectal cancer and dysplasia in inflammatory bowel disease. *World J Gastroenterol* 2008;14:2662–9.
- Cairns SR, Scholefield JH, Steele RJ, *et al*. Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002). *Gut* 2010;59:666–89.
- Laine L, Kaltenbach T, Barkun A, *et al*. SCENIC international consensus statement on surveillance and management of dysplasia in inflammatory bowel disease. *Gastrointest Endosc* 2015;81:489–501.
- Lamb CA, Kennedy NA, Raine T, *et al*. British Society of Gastroenterology consensus guidelines on the management of inflammatory bowel disease in adults. *Gut* 2019;68:s1–106.
- Darie A-M, Sinopoulou V, Ajay V, *et al*. BSG 2024 IBD guidelines protocol (standard operating procedures). *BMJ Open Gastroenterol* 2023;10:e001067.
- Piggott T, Langendam MW, Parmelli E, *et al*. The GIN-McMaster guideline tool extension for the integration of quality improvement and quality assurance in guidelines: a description of the methods for its development. *J Clin Epidemiol* 2023;154:197–203.
- Schünemann H, Brożek J, Guyatt G, *et al*. GRADE handbook for grading quality of evidence and strength of recommendations. *The GRADE Working Group*; 2017.
- Organization WH. WHO Handbook for Guideline Development. World Health Organization, 2014.
- Schünemann HJ, Wiercioch W, Etzeandia I, *et al*. Guidelines 2.0: systematic development of a comprehensive checklist for a successful guideline enterprise. *CMAJ* 2014;186:E123–42.
- McMaster. GIN-mcmaster guideline development checklist, Available: <https://cebgrade.mcmaster.ca/guidelinechecklistonline.html> [Accessed 3 Sep 2023].
- Fretheim A, Schünemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 3. Group composition and consultation process. *Health Res Policy Syst* 2006;4:15.
- Brouwers MC, Kho ME, Browman GP, *et al*. AGREE II: advancing guideline development, reporting and evaluation in health care. *Can Med Assoc J* 2010;182:E839–42.
- Sharma R, Gordon M, Dharamsi S, *et al*. Systematic reviews in medical education: A practical approach: AMEE Guide 94. *Med Teach* 2015;37:108–24.
- Morgano GP, Mbuagbaw L, Santesso N, *et al*. Defining decision thresholds for judgments on health benefits and harms using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Evidence to Decision (EtD) frameworks: a protocol for a randomised methodological study (GRADE-THRESHOLD). *BMJ Open* 2022;12:e053246.
- Zeng L, Brignardello-Petersen R, Guyatt G. When applying GRADE, how do we decide the target of certainty of evidence rating? *Evid Based Mental Health* 2021;24:121–3.
- Alonso-Coello P, Oxman AD, Moberg J, *et al*. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *Gac Sanit* 2018;32:167.
- Guyatt GH, Oxman AD, Schünemann HJ, *et al*. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *J Clin Epidemiol* 2011;64:380–2.
- Schünemann HJ, Neumann I, Hultcrantz M, *et al*. GRADE guidance 35: update on rating imprecision for assessing contextualized certainty of evidence and making decisions. *J Clin Epidemiol* 2022;150:225–42.
- About Jotform, Available: <https://eu.jotform.com/about/> [Accessed 9 Jul 2024].
- Shea BJ, Reeves BC, Wells G, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.
- Sterne JA, Hernán MA, Reeves BC, *et al*. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;i4919.
- Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
- Cumpston M, Li T, Page MJ, *et al*. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. *Cochrane Database Syst Rev* 2019;10:ED000142.
- Whiting P, Rutjes AWS, Reitsma JB, *et al*. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;3:25:1–13.
- Egger M, Smith GD, Schneider M, *et al*. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
- GDT GRADEproGRADEpro gdt: gradepro guideline development tool software.
- Alonso-Coello P, Schünemann HJ, Moberg J, *et al*. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 1: Introduction. *BMJ* 2017;32:i2016.
- Mustafa RA, Garcia CAC, Bhatt M, *et al*. GRADE notes: How to use GRADE when there is “no” evidence? A case study of the expert evidence approach. *J Clin Epidemiol* 2021;137:231–5.