Identifying potential medical causes of fatigue, pain, and urgency in inflammatory bowel disease and optimising medical management of these causes.

The IBD-BOOST OPTIMISE Study

Statistical Analysis Plan

Version: 1.0
Date: 27/04/2023

Person(s) contributing to the analysis plan	
Name(s) and position(s)	Fionn Cléirigh Büttner, trial statistician Thomas Hamborg, senior statistician
Authorisation	
Position	**Chief or principal investigator**
Name	**Christine Norton**
Signature	
Date	27.04.2023
Position	**Senior trial statistician**
Name	**Thomas Hamborg**
Signature	Thomas Hamborg Digitally signed by Thomas Hamborg Date: 2023.04.27 23:54:58 +01'00'
Date	
Position	**Independent statistician**
Name	**Stephanie MacNeill**
Tick once reviewed	✓
Date	12/04/2023

1. Administrative Information

1.1 Trial registration number: ISRCTN15380317

This SAP is based on protocol version 7.0 (date 16/08/2022)

1.2 SAP revision history

Protocol version	Updated SAP version no.	Section number changed	List of changes from previous version/protocol	Author of change	Date
7.0	0.1			Fionn Cléirigh Büttner	17/10/2022
7.0	0.2			Fionn Cléirigh Büttner	04/11/2022
7.0	0.3			Fionn Cléirigh Büttner	22/11/2022
7.0	0.4			Fionn Cléirigh Büttner	01/12/2022
7.0	0.5	6.6 & 6.9	Minor change suggestion by independent statistician: Specify complete case data as primary analysis and multiply imputed data as sensitivity analysis.	Stephanie MacNeill	12/04/2023
7.0	0.6	6.6 & 6.9	Change suggested by independent statistician incorporated. Other relevant track changes and comments removed.	Fionn Cléirigh Büttner	13/04/2023
7.0	0.7	Sections throughout	Incorporate comments by senior statistician (TH) and chief investigator	Fionn Cléirigh Büttner	20/04/2023
7.0	0.8	Sections throughout	Incorporate comments by chief investigator	Fionn Cléirigh Büttner	24/04/2023
7.0	1.0	Sections throughout	Final comments from senior statistician (TH) incorporated.	Fionn Cléirigh Büttner	27/04/2023

*If the SAP has been published, indicate which version.

1.3 Members of the writing committee

The SAP writing committee comprises Fionn Cléirigh Büttner (FCB) and Thomas Hamborg (TH), Christine Norton (CN) provided input. FCB is primarily responsible for writing and implementing the statistical analysis strategy.

1.4 Timing of SAP revisions in relation to unblinding of data/results

No members of the writing committee have, or will have, access to data before the statistical analysis plan is signed off.

1.5 Remit of SAP

The current Statistical Analysis Plan (SAP) outlines detail of statistical analyses and tables to be reported within the principal paper(s) of the IBD-BOOST OPTIMISE non-randomised intervention study (IBD-BOOST WP2c). It details the sample size justification and intended statistical analyses of primary and secondary study aims and outcomes, subgroup analyses, and sensitivity analyses. Unplanned, post hoc analyses will be transparently reported in the study manuscript.

Abbreviations/Glossary of terms

AE	Adverse Events
BCC	Barts Cancer Research UK Centre
CCUK	Crohn's & Colitis UK (registered charity supporting people with IBD)
CD	Crohn's disease
CRF	Case Report Form
CI	Chief Investigator
CV	Curriculum Vitae
FI	Faecal Incontinence
GCP	Good Clinical Practice
GDPR	General Data Protection Regulation
GP	General Practitioner
HRA	Health Research Authority
IBD	Inflammatory Bowel Disease
IRAS	Integrate Research Application System
KCL	King's College London
LNWH	London Northwest Healthcare NHS Trust
NHS	National Health Service
NICE	National Institute for Clinical Excellence
NIHR	National Institute for Health Research
NRES	National Research Ethics Service
PCTU	Pragmatic Clinical Trials Unit (Queen Mary, University of London)
PCfAR	Programme Grant for Applied Research
PI	Principal Investigator
PIS	Patient Information Sheet
PMG	Programme Management Group
PSC	Programme Steering Committee
PPI	Public and Patient Involvement
QMUL	Queen Mary's University of London
RCT	Randomised Control Trial
REC	Research Ethics Committee
SAEs	Serious Adverse Events
SAP	Statistical Analysis Plan
SoA	Statement of Activities
SoE	Schedule of Events
UC	Ulcerative Colitis

2. Introduction

Inflammatory Bowel Disease (IBD) affects 300,000 people in the UK (https://www.crohnsandcolitis.org.uk/about-inflammatory-bowel-disease), causing unpredictable bouts of gut inflammation, with acute illness, diarrhoea, and pain. In remission, many people with IBD live with fatigue, chronic pain, and bowel urgency/incontinence (1). There is no current cure for IBD, which usually starts in childhood or as a young adult. Most previous IBD research has focused on controlling inflammation. However, many people report continuing IBD-related fatigue (41%), abdominal pain (62%) and difficulty with continence (up to 75%) even when IBD is in remission (1-3). These symptoms limit peoples' quality of life and ability to work and socialise. Patients feel these symptoms are not taken seriously by health professionals and report that little help is given (4-6). However, the James Lind Alliance IBD research priority-setting consensus put fatigue, pain, and continence in the top 10 issues that IBD patients and clinicians want to be addressed by research (7).

2.1 The IBD-BOOST programme of research

The current study forms stage three of IBD-BOOST – a National Institute of Health Research (NIHR) Programme Grant for Applied Research (PGfAR) funded programme. The overall aim of the Programme Grant is to improve the quality of life of people with IBD by reducing the burden of IBD-related fatigue, pain, and urgency/incontinence. The current Statistical Analysis Plan is for stage 3 of the programme, a non-randomised intervention study to test a checklist and algorithm for identifying and managing potential medical causes of these IBD-related symptoms.

Stage 1 of the programme involved focus groups and interviews with people with IBD and IBD nurses to inform the development of the checklist and algorithm under investigation in this study.

Stage 2 of the programme involves a large cross-sectional survey of people with IBD to investigate the inter-relationships of these IBD-related symptoms and the proportions wanting support to manage these symptoms. Survey participants who consent to be contacted for further stages of our research, and who attend one of our clinical sites for their routine IBD care, will be invited to participate in the Optimise Study (this stage) and then, if eligible, will later be invited to participate in the RCT.

Stage 3 (this project) of the programme comprises a non-randomised intervention study to test a checklist and clinical management algorithm which we have developed within this programme for identifying and managing the most common medical causes of these IBD-related symptoms. We will then address any medical abnormalities detected.

Stage 4 Participants who have no medical causes identified by the checklist and algorithm, or who have abnormalities and complete this medical optimisation but have continuing symptoms after medical optimisation will then be invited to take part in an RCT of online self-management. The RCT will also be offered to eligible people declining participation in the present study, or who consent to the present study and complete the checklist but who subsequently decline any suggested medical test or intervention.

1.2 Rationale

Symptoms of fatigue, pain, and urgency/incontinence have a major impact on quality of life in people with IBD but have been largely ignored by clinicians and researchers. Our programme,

shaped by the concerns of our patient and clinician stakeholders, focuses on a supported online self-management intervention for these symptoms. The current checklist and algorithm stage will help identify participants who will be suitable for a self-management intervention and ensure that anyone displaying "red-flag" symptoms (indicating an urgent or serious medical issue) is identified for prompt treatment. It is currently unclear how useful it is to investigate these symptoms and whether symptoms will respond to correcting biomedical abnormalities.

1.3 The need for research

It is currently unclear how best to manage these common symptoms of fatigue, pain, and urgency/incontinence in people with inflammatory bowel disease. Many patients do not report these symptoms at all, or if they do are offered little beyond investigation and treatment of active disease. We have found in previous work that many patients do not receive what are considered "standard care" investigations or management for these symptoms (8). Our previous systematic literature reviews (9-12) have identified many potentially reversible causes for these symptoms. Many of these, particularly the psycho-social elements, will be addressed in our online self-management programme which follows on from the current proposal within our programme grant. However, there are also "medical" causes (such as anaemia as a cause of fatigue), which could be addressed before patients enter a self-management programme. However, during extensive consultation with expert clinicians during development of the current checklist and algorithm, there is genuine equipoise over whether patients need these medical causes addressing, whether it is likely to make any difference to symptoms, or whether they should be able to directly access online self-management.

1.4 Development of the checklist and algorithm

Development work for this study has involved consulting a wide variety of gastroenterologists, IBD specialist nurses, patients and others. This work suggests that there is no current "standard of care" for investigating and treating these symptoms. Some clinics do very little while others conduct many different tests, with little consideration of costs or inconvenience to patients. Over an extensive series of interviews and group activities with stakeholders, there is now consensus on the contents of an algorithm for investigating and managing symptoms of fatigue pain and urgency/incontinence, with an accompanying preliminary self-completed checklist. During this development we have focussed on the common evidence-based causes (from our systematic reviews above) and interventions for our three symptoms (fatigue, pain, urgency), which consultees felt were also feasible for implementation in routine clinical practice. We have tried to achieve a balance between feasibility, costs and completeness. There were many other candidate tests, which were either not evidence-based or that patients and clinicians felt were too onerous to be part of routine clinical care for all patients with these symptoms.

3. Background and trial design

Study objectives	1. To determine the proportion of participating IBD patients with fatigue, pain, and/or faecal urgency/incontinence who have pathophysiological contributors that are potentially medically treatable. 2. To describe changes in symptoms three months after initiating algorithm-led management amongst patients with IBD-related symptoms and potentially reversible medical causes. 3. To evaluate changes in health-related quality of life amongst patients with IBD-related symptoms starting algorithm-led management. 4. To ascertain the feasibility of implementing IBD symptom checklist and algorithm in clinical practice. To determine what resources, such as costs, are incurred by the NHS during implementation of the algorithm. 5. To qualitatively evaluate the experience and suggestions of nurses implementing the algorithm.
Study design	Multi-site, non-randomised, intervention study
Setting	Four NHS Trusts: - Barts Health NHS Trust (Royal London Hospital) - Dorset County Hospital NHS Foundation Trust (Dorset County Hospital) - London Northwest University Healthcare NHS Trust (St Mark's Hospital & Northwick Park Hospital) - St Helens and Knowsley Teaching Hospitals NHS Trust (Whiston Hospital or St Helen's Hospital)
Participants	Adults with IBD who have completed the IBD-BOOST survey at stage 2 of the Programme Grant and have indicated on that survey that they would like further support for their symptoms **Number of participants** 200 **Inclusion criteria** • Diagnosis of IBD (including patients with an ileo-anal pouch or stoma) • 18 years and over • Lives in UK and attends one of the IBD-BOOST clinical sites for routine IBD care • Has completed the IBD-BOOST survey (stage 2 of Programme Grant) and indicated that they would like further support to help manage their symptoms • Ability to give informed consent and sufficient command of English to understand study documents and procedures will be assumed from response to the previous survey

	Exclusion criteria • Under 18 years
Interventions	Study participants will complete a self-report symptom checklist and provide a stool sample. A healthcare professional from their clinical site will review checklist responses using an agreed algorithm and offer to initiate additional tests or clinical management as indicated by the algorithm. There is no control group in this non-randomised, intervention study.
Primary outcome measure(s)	**<u>Feasibility outcomes</u>** • Proportion of consenting participants who: (i) completed the symptom checklist (ii) returned the postal faecal calprotectin test • Proportion of participants for whom, after returning both the symptom checklist and faecal calprotectin test, no algorithm-informed action was indicated (and thus could progress to the RCT) • Proportion of participants for whom algorithm-informed action was indicated where the clinical team completed and returned the CRF (to indicate the actions they planned) • Of participants where algorithm-informed actions were indicated, the proportion who completed outcome measures at three months (i.e., attrition rates). • Number and proportion of participants consenting to receive the algorithm-informed intervention but then discontinuing (with reasons if possible). **<u>Clinical outcomes</u>** The proportion of patients with any of the following, detected using the self-report symptom checklist, faecal calprotectin test, or by the nurse/clinician following the algorithm: • "Red flags" on the self-report symptom checklist that require investigation (as a single binary indicator variable). • Active disease (defined as faecal calprotectin $\geq$200 and/or IBD control score $\leq$13). • Participants for whom interventions for fatigue are advised or prescribed (based on detected blood test abnormalities). • Participants for whom interventions for pain are advised or prescribed (based on irritable bowel syndrome or functional dyspepsia, as diagnosed from responses on checklist). • Participants for whom interventions for faecal urgency are advised or prescribed (based on the detection of untreated loose stool).

4. Outcome measures

As the IBD-BOOST Optimise study had to be adapted due to covid-19, with reduction of planned sample size from 500 to 200, the main focus of analysis will be on feasibility outcomes:

- Proportion of consenting participants who:
  (iii) completed the symptom checklist
  (iv) returned the postal faecal calprotectin test
- Proportion of participants for whom, after returning both the symptom checklist and faecal calprotectin test, no algorithm-informed action was indicated (and thus could progress to the RCT)
- Proportion of participants for whom algorithm-informed action was indicated where the clinical team completed and returned the CRF (to indicate the actions they planned)
- Of participants where algorithm-informed actions were indicated, the proportion who completed outcome measures at three months (i.e., attrition rates).
- Number and proportion of participants consenting to receive the algorithm-informed intervention but then discontinuing (with reasons if possible). Discontinuing will be defined as actively notifying study personnel of withdrawal or failing to respond to the OPTIMISE questionnaire at three months if it was indicated.

4.1 Primary clinical outcome measures (upon administration of the self-report symptom checklist and faecal calprotectin test, before delivery of algorithm-informed intervention(s)):

The proportion of patients with any of the following, detected using the self-report checklist, faecal calprotectin test, or by the nurse/clinician after implementing the algorithm:

- "Red flags" on the self-report symptom checklist that require investigation (as a single binary indicator variable).
- Active disease (defined as faecal calprotectin $\geq$200 and/or IBD control score $\leq$13).
- Participants for whom interventions for fatigue are advised or prescribed (based on detected blood test abnormalities).
- Participants for whom interventions for pain are advised or prescribed (based on irritable bowel syndrome or functional dyspepsia, as diagnosed from responses on checklist).
- Participants for whom interventions for faecal urgency are advised or prescribed (based on the detection of untreated loose stool).

4.2 Secondary clinical outcomes

Prior to algorithm-informed intervention(s):

(i) The proportion of participants for whom clinical intervention was indicated by the algorithm.

(ii) The proportion of participants who declined a suggested clinical intervention by the algorithm.

Prior to and following algorithm-informed intervention:

(iii) The following outcome measures will be obtained from the IBD-BOOST stage 2 survey dataset for the prior time point. Outcome assessment will be repeated at three months

after return of the self-report symptom checklist and algorithm for patients who undergo algorithm-informed intervention(s):

- PROMIS Short Form v1.0 – Fatigue 4a [Four-item validated scale to measure fatigue (14)];
- PROMIS Scale v1.0 – Pain Intensity 3a [Three-item validated scale to measure pain (14)];
- PROMIS Scale v1.0 – Gastrointestinal Bowel Incontinence 4a [Four-item validated scale to measure bowel control (15)];
- IBD-Control score [Eight-item self-reported score to measure disease control from the patient's perspective (13)], and;
- EQ-5D-5L (Quality of Life measurement) [Five-item standardised measure of health (16)].

(iv) The economic cost of implementing the algorithm (i.e., cost of clinical tests, intervention(s), and the time required the nurse/clinician to implement algorithm) [Health Economics objective that will be detailed in HEAP and is not being analysed as part of the statistical analysis].

5. Sample size and randomisation

5.1 Initial sample size justification (500 participants)

From the prevalence of symptoms in IBD patients, it is anticipated that at least 50% of 6,300 predicted respondents to the survey (stage 2) will report one or more symptoms of fatigue, pain, or urgency/faecal incontinence (n=3,150 symptomatic individuals). Of symptomatic participants, it is anticipated that approximately one-third will express an interest in further management for these symptoms and attend one of our clinical sites for routine IBD care (~1,000 individuals available as potential recruits for the present study). Of ~1,000 potential recruits, it is anticipated that 50% (500 individuals) will consent to participate in the OPTIMISE study.

5.1a Revise sample size

Due to the impact of the COVID-19 pandemic on study recruitment, with only 4 of the anticipated 20 NHS sites ever opening, it was necessary (with Steering Group, funder, and ethical approval) to reduce the target sample size from 500 participants to 200 participants and convert this to a feasibility study with the focus on feasibility outcomes (see above). With a sample size of 200 participants, the width of the 95% confidence interval (CI) for the proportion of participants with pathophysiological contributors that are potentially medically treatable will be at most 13.9% (normal approximation CI).

5.2 Randomisation procedure

As this is a non-randomised, intervention study, no randomisation procedure is required.

5.3 Blinding of research personnel

The chief investigator, senior statistician, and trial statistician are blinded – that is, did/do/will not access any study data prior to statistical analysis.

6. Analysis methods

6.1 Baseline characteristics

Baseline demographic characteristics and clinical characteristics obtained when conducting the survey (**stage two**) will be presented using descriptive statistics. Categorical variables will be presented as absolute frequencies and percentages. Continuous variables will be presented as means and standard deviations or as medians and inter-quartile ranges, depending on the data distribution. Draft tables are presented in the appendix.

6.2 Data completeness

The number and percentage of study participants with missing data for each variable will be presented.

6.3 Information for flow diagram

The flow of participants through the following phases of the OPTIMISE study will be reported using a flow-diagram modified for the reporting of evaluations of non-randomised, intervention studies:(17)

(i) Enrolment: the number of survey (**stage two**) respondents who (i) are eligible to participate in the OPTIMISE study, (ii) express an interest in receiving further management for their IBD-related symptoms and thus were invited to participate, and (iii) consented and are enrolled in the OPTIMISE study.

(ii) Assignment: the number of study participants who (i) receive the self-report symptom checklist, (ii) complete the self-report symptom checklist, (iii) are administered the faecal calprotectin test, and (iv) return the faecal calprotectin test.

(iii) Allocation & intervention exposure: the numbers of participants (i) for whom the OPTIMISE algorithm indicates that no further action is required (ii) for whom a red flag is identified and no further OPTIMISE intervention is permitted (iii) who are (iv) assigned to (each) intervention, as informed by the OPTIMISE algorithm, and (v) receive (each) intervention.

(iv) Follow-up: the number of participants receiving algorithm-informed management who do and do not complete study follow-up (three months after completing the self-report symptom checklist at baseline).

(v) Analysis: the number of participants included in or excluded from the analyses of primary outcome, outcomes at three-months, and change in patient-reported outcome measures amongst participants who receive algorithm-informed interventions.

A template flow chart is presented in Appendix 1.

6.4 Adherence to treatment

Adherence to algorithm-informed intervention will not be measured in the IBD-BOOST OPTIMISE study.

6.5 General analysis principles

Study participants who express desire for support with IBD-related symptoms and receive the self-report symptom checklist will be included in descriptive baseline analyses. For study participants who receive algorithm-informed intervention(s), the effect size (e.g., mean change in symptom severity scores) will be estimated. A two-sided p-value and 5% alpha level will be used to declare

statistical significance. However, the focus of this analysis is descriptive. Ninety-five percent confidence intervals (95%CIs) will be reported unless stated otherwise.

6.6 Analysis of the primary clinical outcomes

To fulfil the primary study objective, the proportion of study participants with (i) red flags, (ii) active IBD disease, (iii) blood test abnormalities (in participants with fatigue), (iv) irritable bowel syndrome or functional dyspepsia (in participants with pain), and (v) untreated loose stool (in participants with urgency) will be estimated. We will construct 95%CIs around proportion estimates using Wilson's Score Interval to minimize the loss of coverage that typically occurs with normal approximations of the Wald Interval. Complete case analyses (ignoring missing data) will be performed for primary outcomes.

6.7 Checklist and algorithm feasibility

The analysis of feasibility outcomes will be performed in the same way as the analysis of the clinical outcomes – that is, as descriptive statistics with corresponding 95% CIs. All feasibility outcomes are presented in Table 8.

6.8 Analysis of secondary (clinical) outcomes

Prior to algorithm-informed intervention(s):

The proportion of study participants for whom algorithm-informed management is indicated, and the proportion of study participants who decline indicated algorithm-informed management, will be estimated using the same approach taken for the descriptive analysis of the primary outcome(s).

Prior to and following algorithm-informed intervention:

PROMIS scores & Quality of Life scores

Means, standard deviations, and 95%CIs will be calculated for each PROMIS score (i.e., pain T-score, fatigue T-score, and bowel incontinence raw score) and EQ-5D-5L score will be calculated at initial checklist implementation (i.e., baseline) separately by whether intervention is indicated.

Following algorithm-informed intervention:

Change in PROMIS scores & Quality of Life scores

For study participants who receive algorithm-informed interventions for treatable causes of IBD-related symptoms, PROMIS (for fatigue, pain, and faecal urgency/incontinence symptoms), IBD control score, and EQ-5D-5L scores three months after checklist implementation, after adjusting for relevant confounders, will be estimated. A linear regression model will be constructed for each outcome variable listed in appendix Table 6, adjusting for the baseline measure of that outcome variable, IBD diagnosis (e.g., Crohn's disease or other IBD type), and NHS Trust. Linear regression assumptions – for example, linearity, normally distributed error residuals, and residual homoscedasticity, and multi-collinearity – will be assessed. If data are normally distributed, a normal approximation will be used to estimate 95%CIs. If data are not normally distributed, bootstrap CIs will be estimated (using 1,000 bootstrap samples) because sample sizes may not be sufficiently large to invoke the central limit theorem. If regression models fail to converge, the covariate 'NHS Trust' will be removed.

6.9 Interim analyses

Interim analyses of OPTIMISE study data will not be performed.

6.10 Subgroup analyses

A subgroup analysis will be performed that investigates the PROMIS scores (for pain intensity, fatigue, and bowel incontinence) at three months, and change in PROMIS scores for participants who receive the intervention, by disease type (Crohn's disease versus other IBD). Additionally, feasibility outcomes will be investigated by the four NHS Trusts/sites from which participants are recruited.

6.11 Sensitivity analyses

To assess the extent to which study results are affected by missing data, a sensitivity analysis will be performed on imputed data for primary clinical outcomes (23). The proportion of missing values for each variable will be assessed using numerical and graphical summaries. Univariable associations between missing values of each variable and observed values of other variables will be examined to understand how reliably a missing value might be imputed (18). This will be performed by constructing separate logistic regression models after creating a binary indicator variable for each variable with missing values coded as "1" and non-missing values coded as "0". The most applicable missing data mechanism will be informed by clinical knowledge of independent and dependent variables, reasons for missingness, and relationships between missingness and the observed values of collected variables.

Multivariate Imputation using Chained Equations (MICE) will be used to impute missing data under the expectation that both independent and dependent variables will have missing values and the data will not be monotonic missing (19). MICE replaces missing values with a random sample of plausible, imputed values drawn from their predictive distribution (20). First, an 'imputation' step will be performed, which involves constructing an imputation model that replaces missing data with one set of plausible values. Assuming that missing data are 'Missing At Random', the imputation model will specify a conditional distribution for missing values of each variable given the observed values of other variables. This imputation model will repeatedly replace missing values with a random sample of plausible values, creating a completed dataset with each imputation. The number of imputations (and thus completed datasets generated) will mirror the proportion of participants with at least one missing value. For example, 25 complete datasets will be generated if 25% of study participants have at least one missing value (21).

A logistic regression model will be used for missing values of binary variables and a multinomial logistic regression model will be selected for missing values of categorical variables with three or more unordered categories. Missing values of categorical variables with three or more ordered categories will be modelled using ordinal logistic regression and a linear regression model will be specified for continuous variables with missing data. Auxiliary variables – that is, variables that are not included in the intended analysis of imputed variables but are the highly correlated with the imputed variables (or its missingness) – will be included in the imputation model (21). Due to the descriptive nature of specified statistical analyses, no interaction terms or transformations will be included in the imputation model.

Next, an 'estimation' step will be conducted, whereby specified descriptive analyses – as described in sections 6.6 and 6.7 – will be performed separately for each completed dataset that is generated during the imputation step. Finally, a 'pooling' step will be performed, whereby point estimates (e.g., sample means) and measures of precision (e.g., standard deviations) estimated in each dataset will be aggregated using Rubin's Rules to create a final estimate that accounts for between- and within-imputation uncertainty (22).

7. Other analyses, data summaries, and graphs

Other data summaries

Safety analyses

The number and proportion of participants experiencing (i) adverse events (AEs) and (ii) serious adverse events (SAEs) will be tabulated, stratifying by whether or not algorithm-informed interventions were indicated. An AE is defined as any clinical change, disease, or disorder experienced by the participant during their participation in the study, whether or not considered related to the use of the intervention being studied. An AE is defined as serious (i.e., an SAE) if it results in one of the following outcomes:

- A life-threatening AE
- In-patient hospitalisation or prolonged hospitalisation not related to IBD flares, which are expected events
- Persistent or significant disability/incapacity
- A congenital anomaly/birth defect in the offspring of a subject
- Is otherwise considered medically significant by the investigator
- Other medical events requiring intervention to prevent one of the above outcomes.

Graphs

Line graphs will be constructed to visualise group means and 95% confidence intervals at baseline and three months after algorithm-informed interventions for PROMIS pain intensity, fatigue, and bowel incontinence scores.

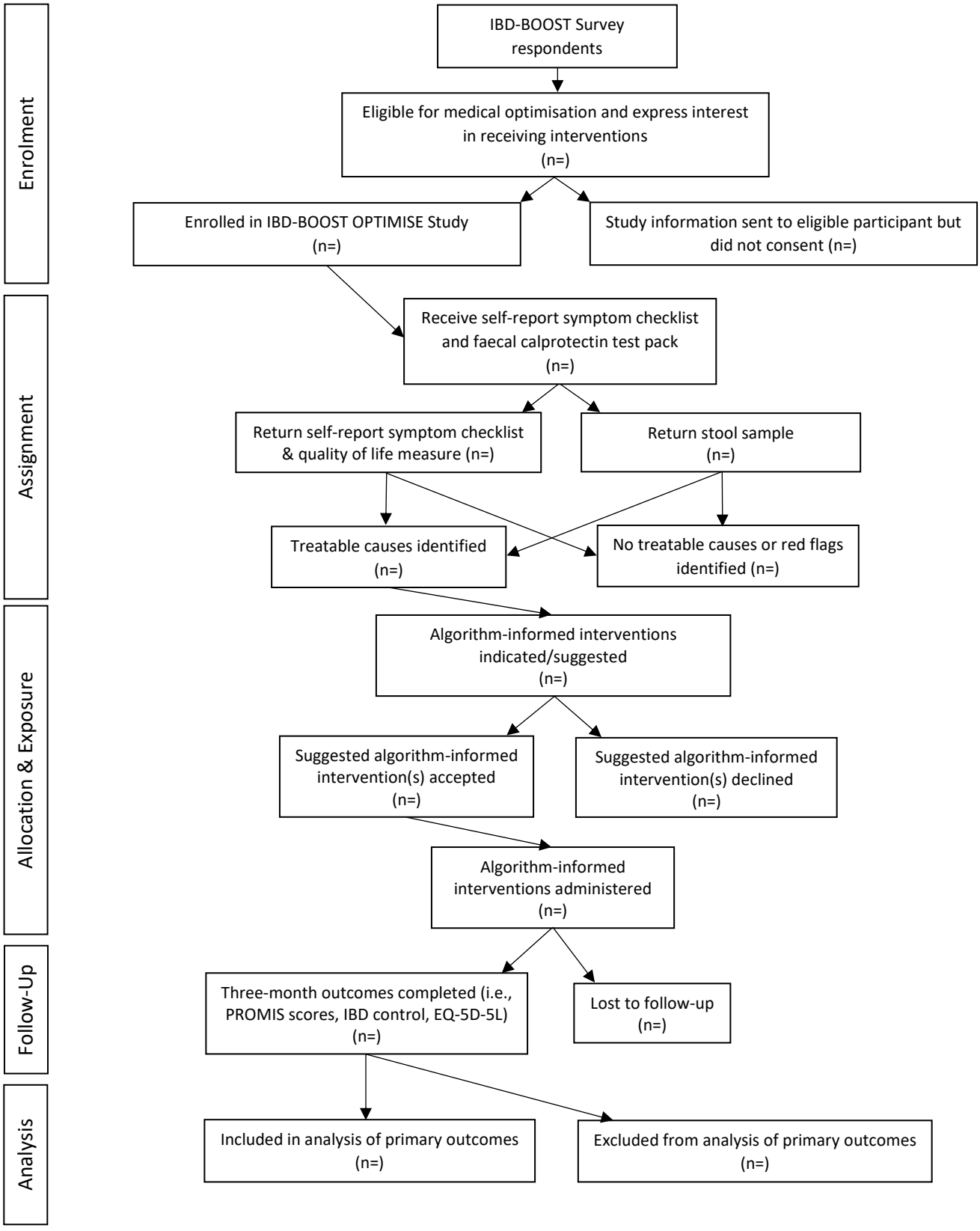
8. References

1. Wilson BS, Lonnfors S, Vermeire S, Greco M, Hommes DW, Bell C, et al. The true impact of IBD: a European Crohn's and Ulcerative Colitis patient life impact survey 2010-11. 2012 2012.
2. Czuber-Dochan W, Ream E, Norton C. Review article: description and management of fatigue in inflammatory bowel disease. Alimentary Pharmacology and Therapeutics. 2013;37:505-16.
3. Norton C, Dibley L, Bassett P. Faecal incontinence in inflammatory bowel disease: associations and effect on quality of life. Journal of Crohn's and Colitis. 2013;doi.org/10.1016/j.crohns.2012.11.004.
4. Czuber-Dochan W, Dibley L, Terry H, Ream E, Norton C. The experience of fatigue in people with inflammatory bowel disease: an exploratory study. Journal of Advanced Nursing. 2012;doi: 10.1111/jan.12060.
5. Czuber-Dochan W, Norton C, Bredin F, Darvell M, Nathan I, Terry H. Healthcare professionals' perceptions of fatigue experienced by people with IBD. Journal of Crohn's & colitis. 2014;doi:.org/10.1016/j.crohns.2014.01.004.
6. Dibley L, Norton C. Experience of fecal incontinence in people with inflammatory bowel disease: self-reported experiences among a community sample. Inflammatory Bowel Diseases. 2013;19(7):1450-62.
7. Lind AJ. Top ten research priorities: inflammatory bowel disease. http://www lindalliance org/top-tens asp [Internet]. 2015 7/8/2015.
8. Czuber-Dochan W, Norton C, Bredin F, Forbes A, Nathan I, Berliner S, et al. Assessing fatigue in patients with inflammatory bowel disease. Gastrointestinal Nursing. 2014;12(8):13-21.
9. Artom M, Czuber-Dochan W, Sturt J, Norton C. Targets for health interventions for inflammatory bowel disease-fatigue. Journal of Crohn's and Colitis. 2016.
10. Sweeney L, Moss-Morris R, Czuber-Dochan W, Meade L, Chumbley G, Norton C. Systematic review: psychosocial factors associated with pain in inflammatory bowel disease. Aliment Pharmacol Ther. 2018;47:715-29.
11. Norton C, Czuber-Dochan W, Artom M, Sweeney L, Hart AL. Systematic review: interventions for abdominal pain management in inflammatory bowel disease. Alimentary Pharmacology and Therapeutics. 2017;doi: 10.1111/apt.14108:1-11.
12. Proudfoot H, Norton C, Artom M, Khoshaba K. Systematic review: targets for health interventions for faecal incontinence in inflammatory bowel disease. Scandinavian Journal of Gastroenterology. 2018;accepted, in press.
13. Bodger K, Ormerod C, Shackcloth D, Harrison M. Development and validation of a rapid, generic measure of disease control from the patient's perspective: the IBD-control questionnaire. Gut. 2013:gutjnl-2013-305600.
14. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The PatientReported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. Journal of clinical epidemiology. 2010;63(11):1179-94.
15. Spiegel BM, Hays RD, Bolus R, Melmed GY, Chang L, Whitman C, et al. Development of the NIH patient-reported outcomes measurement information system (PROMIS) gastrointestinal symptom scales. The American journal of gastroenterology. 2014;109(11):1804.

16. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research. 2011;20(10):1727-36.
17. Des Jarlais DC, Cynthia L, Crepaz N, & the TREND Group. Improving the Reporting Quality of Nonrandomized Evaluations of Behavioural and Public Health Interventions: The Trend Statement. American Journal of Public Health. 2004; 94(3):361-366.
18. Steyerberg EW. Dealing with missing values. In: Steyerberg EW, editor. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York, NY: Springer; 2009. pg 115-37 (Statistics for Biology and Health). Available from: https://doi.org/10.1007/978-0-387-77244-8_7.
19. Jakobsen KC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. 2017; 17: 162-71. https://doi.org/10.1186/s12874-017-0442-1.
20. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res. 201; Mar;20(1):40-9.
21. Austin PC, White IR, Lee DS, van Buuren S. Missing Data in Clinical Research: A Tutorial on Multiple Imputation. Can J Cardiol. 2021; 37: 1322-1331.
22. Rubin DB. Inference and Missing Data. Biometrika. 1976; 63(3):581-92.
23. ThabaneL, Mbuagbaw L, Zhang S, et al. A tutorial on sensitivity analyses in clinical trials: the what, why, when, and how. BMC Medical Research Methodology. 2013; 13: 92-103.

9. Appendices

Appendix 1: Flow-diagram



Appendix 2: Derivations of all outcome measures

Scored Measures

1. **Red flag (binary variable)**

   Red flags will be considered "present" if the patient reports **ALL** of the criteria for pain, vomiting, weight loss, inability to eat, constipation, stool, fever **OR** rectal bleeding listed below.

   A red flag from **pain criteria** will be indicated if **ALL** of the following criteria are present:

   1. The patient has severe stomach pain that may or may not be worse after eating.
   2. The patient has experienced this pain for less than three weeks.
   3. The patient has not spoken to a doctor or nurse about this pain.
   4. The patient has not had any investigations performed by their healthcare team for this pain.
   5. The patient has not had any cause diagnosed for this pain.

   A red flag from **vomiting criteria** will be indicated if **ALL** of the following criteria are listed:

   1. The patient has regular or persistent vomiting.
   2. The patient has experienced vomiting for less than two weeks.
   3. The patient has not spoken to a doctor or nurse about this vomiting.
   4. The patient has not had any investigations performed by their healthcare team for this vomiting.
   5. The patient has not had any cause diagnosed for this vomiting.

   A red flag from **weight loss criteria** will be indicated if **ALL** of the following criteria are listed:

   1. The patient has recently experienced unintentional, rapid weight loss without trying to lose weight.
   2. The patient has experienced this weight loss for less than two weeks.
   3. The patient has lost more than 5 kilograms (10 pounds).
   4. The patient has not spoken to a doctor or nurse about this weight loss.
   5. The patient has not had any investigations performed by their healthcare team for this weight loss.
   6. The patient has not had any cause diagnosed for this weight loss.

   A red flag from **inability to eat** criteria will be indicated if **ALL** of the following criteria are listed:

   1. The patient is unable to eat anything except a very soft diet because solid food causes pain.
   2. The patient has experienced this inability to eat for less than two weeks.
   3. The patient has not spoken to a doctor or nurse about this inability to eat.

4. The patient has not had any investigations performed by their healthcare team for this inability to eat.
5. The patient has not had any cause diagnosed for this weight loss.

A red flag from **constipation** criteria will be indicated if **ALL** of the following criteria are listed:

1. The patient has constipation plus inability to pass gas and a very distended abdomen.
2. The patient has experienced this constipation for less than two weeks.
3. The patient has not spoken to a doctor or nurse about this constipation.
4. The patient has not had any investigations performed by their healthcare team for this constipation.
5. The patient has not had any cause diagnosed for this constipation.

A red flag from **stool** criteria will be indicated if **ALL** of the following criteria are listed:

1. The patient has had new symptoms over the past two weeks that comprise passing dark, black, treacle-like stools.
2. The patient has experienced this black stool for less than two weeks.
3. The patient has not spoken to a doctor or nurse about these stools.
4. The patient has not had any investigations performed by their healthcare team for these black stools.
5. The patient has not had any cause diagnosed for these black stools.

A red flag from **fever** criteria will be indicated if **ALL** of the following criteria are listed:

1. The patient experiences fevers not due to an obvious reason such as the 'flu' or symptoms of COVID-19.
2. The patient has experienced these fevers for less than two weeks.
3. The patient has not spoken to a doctor or nurse about these fevers.
4. The patient has not had any investigations performed by their healthcare team for these fevers.
5. The patient has not had any cause diagnosed for these fevers.

A red flag from **rectal bleeding** criteria will be indicated if **ALL** of the following criteria are listed:

1. The patient experiences more than just a few drops of rectal bleeding.
2. The patient has experienced this rectal bleeding for less than two weeks.
3. The patient has not spoken to a doctor or nurse about this rectal bleeding.
4. The patient has not had any investigations performed by their healthcare team for this rectal bleeding.
5. The patient has not had any cause diagnosed for this rectal bleeding.

2. PROMIS fatigue score

The fatigue score is based on responses to four questions: items fatigue3-fatigue9 in the RSD. Each question has 5 responses, coded 1 (i.e., "Not at all") to 5 (i.e., "Very much"). The raw score is the sum of the four responses, and ranges from 4 to 20. The score is invalid if any answers are missing. The raw score can then be converted to a T-score based on the table below

PROMIS Adult v1.0 Fatigue 7a *Short Form Conversion Table*		
Raw Score	**T-score**	**SE**
4	33.7	4.9
5	39.7	3.1
6	43.1	2.7
7	46.0	2.6
8	48.6	2.5
9	51.0	2.5
10	53.1	2.4
11	55.1	2.4
12	57.0	2.3
13	58.8	2.3
14	60.7	2.3
15	62.7	2.4
16	64.6	2.4
17	66.7	2.4
18	69.0	2.5
19	71.6	2.7
20	75.8	3.9
*SE = Standard Error on T-score metric		

Interpretation note: A higher PROMIS T-score represents more of the concept being measured. For negatively worded concepts like fatigue, a T-score of 60 is one SD worse than average. By comparison, a fatigue T-score of 40 is one SD better than average.

3. PROMIS pain intensity score

The pain intensity score is based on responses to three questions: items pain4, pain5, and pain6 in the Requirements Specification Document (v4.0, 24.08.2021). Each question has five responses, coded 1 (i.e., "Had no pain") to (i.e., "Very severe pain") 5. The raw score is the sum of the 3 responses, and ranges from 3 to 15. The score is invalid if any answers are missing. The raw score can then be converted to a T-score based on the table below.

PROMIS Adult v1.0 Pain Intensity 3a *Short Form Conversion Table*		
Raw Score	**T-score**	**SE**
3	36.3	5.4
4	43.1	3.9
5	47.5	3.7
6	51.4	3.8
7	54.8	3.9

8	58.5	3.9
9	61.9	3.8
10	64.9	3.9
11	68.4	4.1
12	72.0	4.2
13	75.1	4.8
14	77.8	5.0
15	81.8	4.2
*SE = Standard Error on T-score metric		

Interpretation note: *A higher PROMIS T-score represents more of the concept being measured.* For negatively-worded concepts like Pain Intensity, a T-score of 60 is one SD worse than average. By comparison, a Pain Intensity T-score of 40 is one SD better than average.

4. PROMIS bowel incontinence score

The bowel incontinence score is based on responses to four questions: items urgency3-urgency6 in the RSD. Each question has five responses, coded 1 (i.e., "No days"/ "Never") to 5 (i.e., "6-7 days"/"Always"). The raw score is the sum of the five responses, and ranges from 4 to 20. The score is invalid if any answers are missing. The scores are not based on item response theory models; therefore, there is no T-score. The raw score is then used for statistical analysis. Higher scores indicate increased incontinence. No items can be skipped and there are no screening questions.

5. IBD control score

The IBD-Control-8 sub-score is the sum of RSD items ibdcontrol1, ibdcontrol2, and ibdcontrol4-ibdcontrol9. The range of scores is 0-16, with 0 indicating worse control and 16 indicating best control. Responses to each individual item comprise: zero points for the least favourable reply; one point for an intermediate/indeterminate reply, and; two points for the most favourable reply. Note that N/A, denoted 999, is actually counted as 1 for item ibdcontrol2. If ibdcontrol2 has the value 999, treat it as having the value 1 to sum the score. The IBD-Control-VAS score indicates self-reported overall level of control and ranges from 0-10, with 0 indicating worst control.

6. EQ-5D-5L

Each of the five dimensions, mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, comprising the EQ-5D descriptive system is divided into five levels of perceived problems:

LEVEL 1: indicating no problem
LEVEL 2: indicating slight problems
LEVEL 3: indicating moderate problems
LEVEL 4: indicating severe problems
LEVEL 5: indicating unable to/extreme problems

A unique health state is defined by combining one level from each of the five dimensions. Levels 1-5 are coded as 1-5, respectively, and combining a level for each dimension gives a five-digit health state. These health states are then converted to an indexed value using the EQ-VT for England. The Index value set calculators are found on the Euroqol website and have been saved in the BOOST BCC folder as a Stata do file (Z:\STATISTICIANS\IBD-BOOST\EQ-5D-

5L index value calculator). The EQ-VAS score is continuous and ranges from 1-100, with 100 indicating “the best health you can imagine” and 0 indicating “the worst health you can imagine.” This score shows the patient’s perceived overall health and requires no further derivation.

Requirements Specification Document Primary Clinical Outcomes

Clinical outcome – Abnormalities detected on blood test in people with fatigue
- Optimisation CRF: **130. fatiguenew D1.1**. Were new tests ordered for causes of fatigue? If that answer is YES and then **198. ftginterven E1.1.** Were interventions for fatigue advised or prescribed? If that is YES that means there was something abnormal.

Clinical outcome – Irritable Bowel Syndrome or functional dyspepsia indicated in people with pain (by responses on checklist)
- Optimisation CRF: **266. paininterven F1.1.** Were interventions for pain advised or prescribed? If that is a YES that means they had IBS or functional dyspepsia pain. We use a scoring system for these which I can send if helpful using the answers on the checklist. Redcap does this automatically. You can see this on the RSD section 3.7 questions 4 and 5.

Clinical outcome - Untreated loose stool detected in people with faecal urgency.
- Optimisation CRF - **280. urgencyadv G1.1. Were interventions for urgency advised or prescribed?** If this is a YES it means they had untreated urgency.

Appendix 3: Descriptive analysis tables

Table 1: Number and percentage of OPTIMISE study participants by site

Site	*N* approached	Response rate (%)
Barts Health NHS Trust (Royal London Hospital)		
Dorset County Hospital NHS Foundation Trust (Dorset County Hospital)		
London Northwest University Hospital NHS Trust (St Mark's Hospital or Northwick Park Hospital)		
St Helens and Knowsley Teaching Hospitals NHS Trust (Whiston Hospital or St Helen's Hospital)		
TOTAL		

Table 2. Demographic and lifestyle characteristics at baseline

Characteristic	Total	Algorithm-informed intervention(s) indicated	Algorithm-informed intervention(s) not indicated
Gender, n (%)			
Male			
Female			
Prefer to self-define			
Prefer not to say			
Missing			
Age (years), mean (sd)			
Ethnicity, n (%)			
Asian			
Black			
Mixed			
Other			
White			
Prefer not to say			
Missing			
Employment, n (%)			
Full time			
Part time			
Student			
Retired			
Unemployed			
Self-employed			
Homemaker			
Unemployed due to illness/disability			
Missing			
Education, n (%)			
No formal education			
Secondary school			
Sixth form			
Further education			
Higher education			
Missing			
Relationship Status, n (%)			
Married/ civil partnership			
Living with partner			
Widowed			
Divorced/separated			
Single			
With partner, not living together			

Missing	
Smoker, n (%)	
Never	
Ex	
Current	
Missing	
Alcohol (units/week) – mean (sd)	
Missing	

Table 3: Clinical characteristics at baseline

Characteristic	Total	Algorithm-informed intervention(s) indicated	Algorithm-informed intervention(s) not indicated
BMI (kg/m^2) – mean (sd)			
Underweight, n(%)			
Healthy weight, n(%)			
Overweight, n(%)			
Missing			
Operation, n(%)			
Missing			
Stoma, n(%)			
Missing			
Pouch, n(%)			
Missing			
Fistula, n(%)			
Missing			
Biologic medication, n(%)			
Missing			
Non-biologic medications, n(%)			
Missing			
Mental health conditions, n(%)			
Missing			
Physical health conditions, n(%)			
Missing			
Pregnant, n(%)			
Missing			
Depression, n(%)			
None			
Mild			
Moderate			
Moderately severe			
Severe			
Missing			
Anxiety, n(%)			
None			
Mild			
Moderate			
Severe			
Missing			

Table 4: Number and percentage of participants with primary outcomes (complete case analysis)

Outcome	n (% of non-missing)
At least one red flag on self-report checklist	n (%)
Active disease*	n (%)
Participants advised or prescribed interventions for **fatigue**	n (%)
Participants advised or prescribed interventions for pain	n (%)
Participants advised or prescribed interventions for faecal urgency	n (%)
Clinical intervention indicated by algorithm	n (%)
Algorithm-informed intervention declined	n (%)

*(faecal calprotectin test ≥ 200 and/or IBD control score ≤13)

Table 5: PROMIS scores for pain, fatigue, and bowel incontinence at baseline and three months after checklist implementation in participants for whom algorithm-informed interventions were, and were not, indicated

Outcome	Intervention indicated mean (SD)	No intervention indicated mean (SD)
PROMIS – Fatigue (T-Score)		
Baseline		
Three months		
PROMIS – Pain Intensity (T-Score)		
Baseline		
Three months		
PROMIS – Bowel Incontinence (Raw Score)		
Baseline		
Three months		
IBD-Control score (Raw Score)		
Baseline		
Three months		
IBD-Control VAS (Raw Score)		
Baseline		
Three months		
EQ-5D-5L		
Baseline		
Three months		
EQ-5D-5L VAS		
Baseline		
Three months		

Table 6: Number (n) and percentage (%) of missing data for PROMIS pain intensity, fatigue, and bowel incontinence scores at baseline and three months after checklist implementation in participants for whom algorithm-informed interventions were, and were not, indicated

Outcome	Intervention indicated n (%)	No intervention indicated n (%)
PROMIS – Fatigue (T-Score)		
Baseline		
Three months		
PROMIS – Pain Intensity (T-Score)		
Baseline		
Three months		
PROMIS – Bowel Incontinence (Raw Score)		
Baseline		
Three months		
IBD-Control score (Raw Score)		
Baseline		
Three months		
IBD-Control VAS (Raw Score)		
Baseline		
Three months		
EQ-5D-5L		
Baseline		
Three months		
EQ-5D-5L VAS		
Baseline		
Three months		

Table 7. Adjusted* mean PROMIS (pain, fatigue, and bowel incontinence) and EQ-5D-5L scores at three months after algorithm-informed interventions were indicated

Outcome	Intervention indicated mean (SD)	Missing n (%)
PROMIS – Fatigue (T-Score)*		
PROMIS – Pain Intensity (T-Score)*		
PROMIS – Bowel Incontinence (Raw Score)*		
IBD-Control score (Raw Score)*		
IBD-Control VAS (Raw Score)*		
EQ-5D-5L*		
EQ-5D-5L VAS*		

*adjusted for confounders: baseline measure for each outcome variable, inflammatory bowel disease diagnosis, and NHS trust

Table 8 Total eligible sample size (N), and the number (n) and percentage (%) of participants completing outcomes on the feasibility of the self-report checklist and management algorithm

Outcome	N	n (%)
Participants consenting to receive algorithm-informed intervention(s) but discontinuing		
Completion rate of outcome measures		
PROMIS – Fatigue		
PROMIS – Pain Intensity		
PROMIS – Bowel Incontinence		
IBD-Control score		
EQ-5D-5L		
Proportion of consenting participants who: (i) completed the symptom checklist (ii) returned the postal faecal calprotectin test		
Proportion of participants for whom, after returning both the symptom checklist and faecal calprotectin test, no algorithm-informed action was indicated (and thus could progress to the RCT)		
Proportion of participants for whom algorithm-informed action was indicated where the clinical team completed and returned the CRF (to indicate the actions they planned)		
Of participants where algorithm-informed actions were indicated, the proportion who completed outcome measures at three months: - PROMIS fatigue score - PROMIS pain intensity score - PROMIS bowel incontinence score - IBD control score - EQ-5D-5L score		

Table 9: Subgroup analysis - adjusted mean PROMIS (pain, fatigue, and bowel incontinence) and EQ-5D-5L scores at three months after algorithm-informed interventions were indicated, stratified by Inflammatory Bowell Disease type

Outcome	Algorithm-informed intervention indicated	
	Crohn's disease (n=X)	Other IBD (n=Y)
PROMIS – Fatigue (T-Score)*	mean (SD)	mean (SD)
PROMIS – Pain Intensity (T-Score)*	mean (SD)	mean (SD)
PROMIS – Bowel Incontinence (Raw Score)*	mean (SD)	mean (SD)
IBD-Control score (Raw Score)*	mean (SD)	mean (SD)
IBD-Control VAS (Raw Score)*	mean (SD)	mean (SD)
EQ-5D-5L*	mean (SD)	mean (SD)

*adjusted for confounders: baseline measure for each outcome variable, inflammatory bowel disease diagnosis, and NHS trust

Table 10: Sensitivity analysis: multiple imputation of primary outcomes

Outcome	n (%) (complete case data)	n (%) (multiply imputed data)
At least one red flag on self-report checklist		
Active disease*		
Participants advised or prescribed interventions for fatigue		
Participants advised or prescribed interventions for pain		
Participants advised or prescribed interventions for urgency		