Evidence Assessment and Analysis Report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence Diagnostics Assessment Programme – Protocol

NIHR Reference No. 135637

Final amended version, 05/12/2022

Title of the project

Quantitative faecal immunochemical tests to guide colorectal cancer pathway referral in primary care

Name of External Assessment Group (EAG) and project lead

School of Health and Related Research Technology Assessment Group (ScHARR-TAG), The University of Sheffield

Project and Systematic Review Lead

Sue Harnan

Senior Research Fellow

School of Health and Related Research

[redacted]

Tel [redacted]

[redacted]

Health Economics Lead

Aline Navega Biz

Research Associate in Economic Modelling

School of Health and Related Research

[redacted]

[redacted]

[redacted]

[redacted]

Tel [redacted]

Plain English Summary

Bowel (colorectal) cancer is the fourth most common cancer in the UK. There is a better chance that bowel cancer can be treated successfully if it is found early. When a patient visits their GP with symptoms, the doctor has to decide whether they think the patient might have bowel cancer or not.

Currently, if the symptoms are “low-risk” and suggest bowel cancer is unlikely, the doctor will ask the person for a faeces sample to conduct a laboratory test called a quantitative faecal immunochemical test (or FIT for short). FIT looks for very small amounts of blood in the faeces, which might be a sign of cancer in the bowel. If their symptoms are “high-risk” and suggest bowel cancer is more likely, the doctor might refer them to be seen by a specialist within two weeks. If the symptoms suggest that another bowel condition is likely, such as inflammatory bowel disease, the doctor might refer the patient for a “non-urgent” specialist appointment, which should happen within 18 weeks.

Specialist appointments may include a test in hospital called a colonoscopy where a special camera is inserted through the anus to look inside the bowel, or CT colonography where images of the bowel are taken from outside the body. Some patients may be offered different tests instead, based on their age and other conditions they may have.

Colonoscopies are good at detecting bowel cancer and other bowel diseases, but are unpleasant, carry a risk of damaging the bowel and there is limited capacity within the NHS to do them. Doctors have noticed that most of the people with high-risk symptoms who are sent for a specialist appointment do not have cancer and did not need an urgent colonoscopy. For these reasons, it is important that only the people most likely to have colorectal cancer are selected to be seen within two weeks. This could then mean that they get the treatment they need quickly, and people with other bowel conditions can also be seen more quickly.

This project will look to see if FIT could be used to identify which people with signs or

symptoms of CRC are most likely to have cancer and need to have an urgent colonoscopy. People who have a positive test would go to hospital for an urgent colonoscopy or CT colonography to see if they have bowel cancer. People who have a negative test would not go to hospital for further tests straight away, but their doctor may refer them for non-urgent appointment. Some may be treated in primary care depending on their symptoms and may go for tests later on if they still have symptoms, or if new symptoms start.

For FIT to be safe and good value for money, it will need to be very good at correctly showing who does and does not have bowel cancer. We will look for research about how many people were correctly and incorrectly identified by the test. We will use mathematical models to determine whether the test is good value for money for the NHS. We will also look at how the threshold (value) used to define a positive test impacts on how many people are sent for colonoscopies, how many people with cancer are missed and have a delay to their diagnosis, and how this impacts on the waiting times for people sent for a non-urgent appointment.

1. Decision problem

1.1 Purpose of the decision to be made

Colorectal cancer (CRC) is the fourth most common form of cancer in the UK. Approximately 42,000 new cases of CRC are diagnosed each year, resulting in around 16,800 CRC-related deaths annually.[1] Quantitative faecal immunochemical tests (FIT) are designed to detect occult (small amounts) of blood in stool samples (faecal haemoglobin) using antibodies specific to human haemoglobin. The UK bowel cancer screening programme currently uses the FIT test for people aged 60 to 74 without any signs or symptoms of cancer. An expansion programme started in 2021 to extend testing to those aged 50-59.

In addition, these tests were recommended by NICE in 2017 (Diagnostic Guidance 30, DG30)[2] to guide referrals for CRC in primary care for low-risk symptomatic patients, i.e. people without rectal bleeding who have unexplained symptoms but do not meet the criteria for a suspected cancer pathway referral, according to NICE Guideline 12 (NG12).[3] Patients who do meet the NG12 criteria are currently referred straight to the urgent two week wait (2WW) suspected CRC referral pathway. Urgent 2WW suspected CRC specialist appointments very often involve full colonic imaging usually with a colonoscopy (COL) but also could involve CT Colonography (CTC). However, since the COVID pandemic and as a result of increasing limitations on these tests, many clinical services in the UK started using FIT in patients presenting with high risk symptoms to reduce referrals.[4, 5]

Symptom-based criteria for referral resulted in an increase in the number of 2WW referrals but there has not been a corresponding increase in the proportion of patients that are investigated who have cancer.[6] Indeed in 2018, of 392,588 referrals made with suspected cancer on the two week wait pathway in England only 13,168 (3.3%) had a cancer. In addition, in August 2022, 28% of people seen by a specialist for suspected CRC were not seen within 2 weeks of urgent referral, and 53% did not have a diagnosis within 28 days (NHS cancer waiting times, August 2022). Of 15,053 people treated for lower gastrointestinal cancer in 2020-21 under a suspected cancer pathway referral, only 50.6% received treatment within 62 days following an urgent GP referral

(compared with an operational standard of 85%).

NICE also heard that wait times for the non-urgent referrals are extremely long in some areas. Amongst patients who present in primary care with symptoms of CRC, non-urgent referrals, usually with an 18 week wait (18WW) target, may be made for patients who do not meet the criteria for a 2WW referral, but for whom there is clinical concern. This may be because the GP suspects another bowel pathology could be present, such as inflammatory bowel disease (IBD, a term used to describe Crohn's disease (CD) and ulcerative colitis (UC)). A delay in diagnosis for these patients could result in worse quality of life and patient outcomes.

The reasons for the increased waiting list times for COL are unclear and may be due to a backlog that accumulated during the coronavirus pandemic, and/or due to referrals exceeding capacity.

NICE heard via consultation with stakeholders and the NHS that the current symptom-based referral pathway, using the NG12 and DG30 criteria, is difficult for GPs to implement. The Association of Coloproctology of Great Britain and Ireland (ACPGBI) / British Society of Gastroenterology (BSG) guideline[7] and the meta-analysis that informed the guidelines[8] also found that there is no clinically significant difference in test accuracy when FIT is used in patients presenting with DG30 and NG12 symptoms as well as those presenting with certain individual symptoms (rectal bleeding, iron deficiency anaemia and abdominal pain), though this guideline did not consider the impact of disease prevalence by symptoms on cost-effectiveness.

There is evidence that FIT may be a better predictor of CRC risk in patients than symptoms alone and could result in fewer referrals of people without CRC to the urgent 2WW suspected CRC pathway. Therefore, triage with FIT could mean that people who are unlikely to have CRC may avoid COL and its associated adverse events (for example, bleeding, perforation and death), and those that are likely to have CRC can be prioritised more effectively[9] leading to a reduction in time to diagnosis. This may also release COL capacity to allow people on non-urgent referral pathways to be seen more

quickly. This will be dependent in part on the threshold used to define a positive test for the symptomatic patients presenting on the 2WW pathway.

The medical technologies topic oversight group identified FIT as an adjunct to clinical assessment in guiding referral for people with high-risk symptoms in primary care as suitable for guidance development by the Diagnostics Assessment Programme (DAP) on the basis of a briefing note. The topic completed scoping in April 2020 but was paused due to changes in clinical pathways due to the COVID-19 pandemic. Following exceptional surveillance of suspected cancer: recognition and referral (NICE guideline NG12) and quantitative FIT to guide referral for CRC in primary care (NICE diagnostics guidance 30), it was decided to resume the topic but rescope to take into account the changes to clinical practice.

As a result of the rescoping exercise, and of the scoping workshop on the 11th of October 2022 and the assessment subgroup meeting on the 2nd of November 2022, the need to identify the optimal way to use FIT to reduce the number of people without significant bowel pathology who are referred to the suspected CRC pathway, taking into consideration the threshold used to define a positive test, and the potential COL capacity constraints for urgent and non-urgent referrals, was identified as an objective of this assessment.

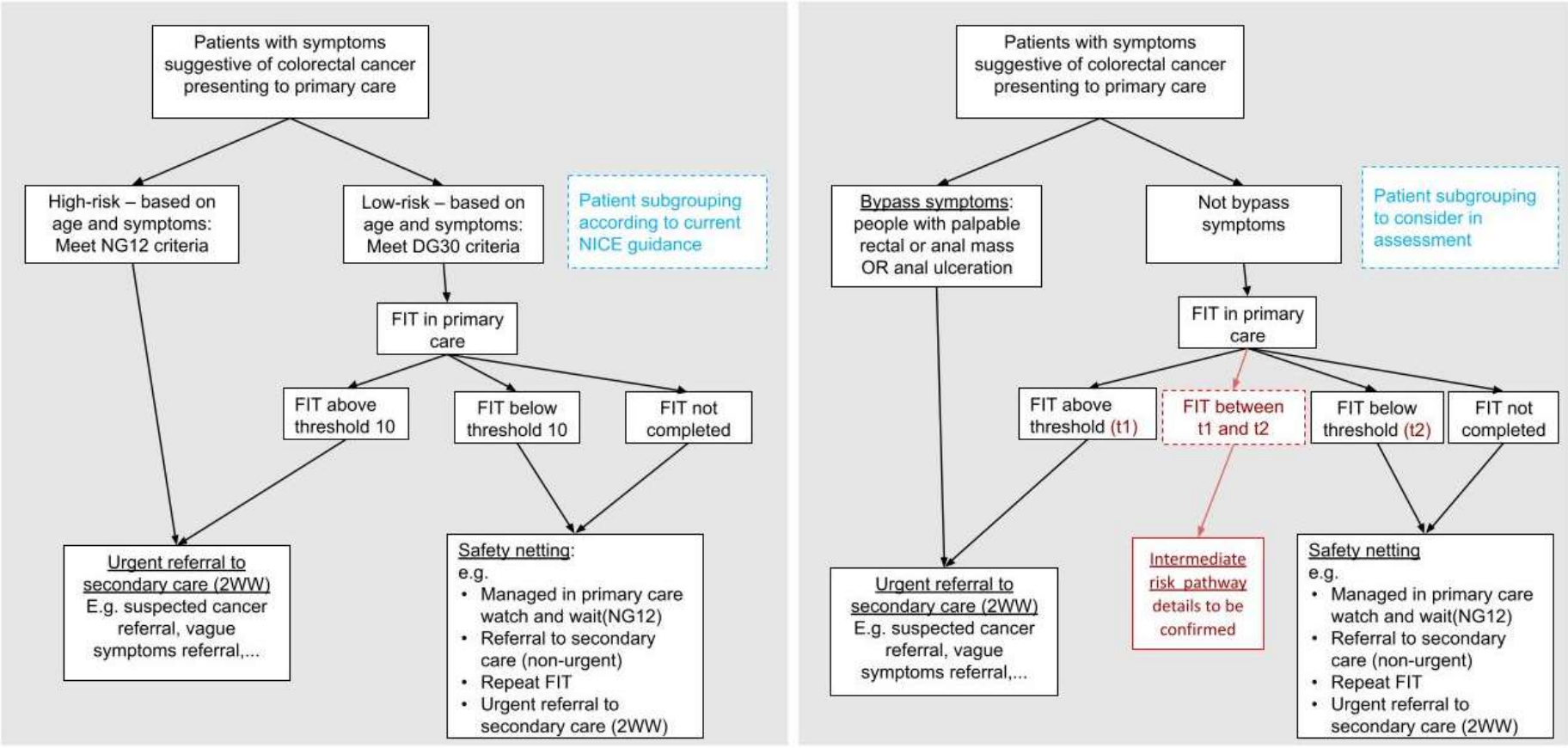
1.2 Place of the intervention in the treatment pathway

This assessment will consider the use of FIT in people presenting to primary care with gastrointestinal symptoms indicating a risk of CRC (excluding those with rectal or anal mass, or anal ulceration who will go straight to urgent 2WW suspected CRC referral). The treatment pathway and proposed position for FIT are shown in Figure 1.

1.2.1 National guideline 12 (NG12) high-risk and Diagnostic guideline 30 (DG30) low-risk patients

NG12 describes the diagnostic pathway for patients presenting to primary care with symptoms suggestive of CRC.[3] Within this guideline, patients who are considered to

Figure 1: The diagnostic pathway and proposed pathway for patients presenting to primary care with symptoms of colorectal cancer. Based on NG12,[3] DG30[2] and the ACPGBI/BSG guideline.[7]



Note: Intervention 2 (using two FIT thresholds to determine management) includes the pathway shown in red.
2WW, urgent two week wait; FIT, faecal immunochemical tests; DG30, Diagnostic Guidance 30; NG12, NICE Guideline 12, t1, threshold 1; t2, threshold 2

be at high-risk for CRC should be referred to secondary care with an urgent 2WW suspected CRC referral. In secondary care, a specialist may order subsequent tests such as COL or computed tomography colonography (CTC). In some parts of the country, the referral may be made directly to one of these imaging tests. NG12[3] states:

Refer people using a suspected cancer pathway referral (for an appointment within 2 weeks) for CRC if:

- they are aged 40 and over with unexplained weight loss and abdominal pain or
- they are aged 50 and over with unexplained rectal bleeding or
- they are aged 60 and over with:
    - iron-deficiency anaemia or
    - changes in their bowel habit, or
- tests show occult blood in their faeces.

A suspected cancer referral (for an appointment within 2 weeks) should also be considered for:

- People with a rectal or abdominal mass
- Adults aged under 50 with rectal bleeding and any of the following unexplained symptoms or findings:
    - Abdominal pain
    - Change in bowel habit
    - Weight loss
    - Iron-deficiency anaemia.

In July 2017, NG12[3] was partially updated by DG30.[2] In this update, the guaiac faecal occult blood test was replaced with FIT, and hence DG30[2] recommends FIT for suspected CRC in people without rectal bleeding who have unexplained symptoms but do not meet the criteria for a suspected cancer pathway referral, that is, they are at low-risk of CRC. These patients include:

- People aged 50 and over with unexplained:
    - abdominal pain or

    - weight loss, or
- People aged under 60 with
    - changes in their bowel habit or
    - iron-deficiency anaemia or
- People aged 60 and over and have anaemia even in the absence of iron deficiency

Patients testing positive by FIT are referred on to the 2WW suspected CRC pathway. What happens in secondary care following referral is thought to be heterogeneous across England; it may be to a specialist who will order further tests (COL, CTC, or other tests as they see fit) or may be a direct referral by a GP to COL or CTC. The choice of imaging test may be dependent on local practice guidelines or age and comorbidities that contraindicate COL. CTC may be necessary where COL fails. During COL, a biopsy may be taken for histological confirmation, unless this is contraindicated (e.g., blood clotting disorders).

Patients testing negative by FIT are followed up in primary care. This should include "safety netting" as described for all cancer pathways in NG12, to avoid missing disease (cancer or otherwise) in people with negative FIT results (see Section 1.9.5).[3] The review for DG30 found that FIT had high sensitivity (few false negatives), so could be used to safely rule out CRC. However, safety netting in NG12 includes an awareness of the possibility of false-negatives, and re-testing either after a period of time or upon the emergence of new symptoms, or the recurrence, persistence or worsening of existing symptoms.[3]

1.2.2 Speciality guide during the height of the coronavirus pandemic

In November 2020, NICE issued a speciality guide for patient management during the coronavirus pandemic on triaging patients with lower gastrointestinal symptoms, which was supported by the BSG. The advice was to continue to refer according to NG12, but that the use of FIT could be used to help clinicians prioritise referrals. People with more than 100 micrograms Hb/g (µg Hb/g) and no COL within the last 3 years, or who had symptoms considered by a specialist GI surgeon/gastroenterologist to warrant urgent

investigation, would be referred for urgent COL or computerised tomography (CT) which could be CTC or plain CT. People with between 10 and 100 µg Hb/g, or people with more than 100 µg Hb/g who have had a COL requiring no further investigation in the last 3 years, would be referred for prioritised COL or colonic imaging (CTC, plain CT or colon capsule endoscopy, where a small capsule containing a camera is swallowed in order to image the digestive tract). People with less than 10 µg Hb/g would be managed using a safety netting process, which may include strategies for diagnosing other gastrointestinal conditions, and further monitoring for colorectal or other types of cancer.

1.2.3 ACPGBI/BSG guideline and NHS England letter

In 2022, the ACPGBI and the BSG published guidance on FIT in patients with signs or symptoms of suspected CRC (ACPGBI/BSG guidance).[7] This guidance was based on a systematic review of the available evidence, expert opinion and agreed by consensus. Economic evaluation was not conducted. In October 2022, NHS England published letters[10, 11] endorsing the use of the ACPGBI and BSG guidance on FIT in primary care, stating it should be implemented in full.

The ACPGBI/BSG guideline recommends that FIT should be used in primary care to identify people with clinical features of CRC for referral for urgent investigation, using a threshold of 10 µg haemoglobin per gram faeces (µg Hb/g). Those with a FIT of fHb ≥10µg Hb/g should be referred on the urgent 2WW suspected CRC pathway in secondary care. Those not meeting these criteria and with no ongoing clinical concerns can be managed in primary care or referred on an alternative pathway.

The ACPGBI/BSG guideline notes that FIT should not be the sole arbiter of referral. Patients without symptoms were not considered in the guideline and should not be referred on the basis of a positive FIT, except within the context of the national screening programme. Patients with negative FIT should not be excluded from referral; where FIT is <10µg Hb/g but there are persistent and unexplained symptoms which concern the GP, the patient should be referred to secondary care for evaluation. This referral may be to routine or urgent pathways, but not necessarily to the CRC pathway.

Those with abdominal mass should be referred and a FIT ordered at the same time for use in secondary care. Those with anal/rectal mass or anal ulceration should be referred on the urgent 2WW suspected CRC pathway without a FIT.

Experts in secondary care said (during the scoping workshop) that FIT results are often used to inform the choice of further investigation (e.g., COL, CTC, or colon capsule endoscopy) when COL capacity is limited.

The NHS England letter also contains recommendations on safety netting for people with negative FIT results. This is discussed in more detail in section 1.9.5.

The ACPGBI/BSG guideline also includes recommendations for patients who fail to complete their FIT test. This includes informing the patient that their clinical assessment is incomplete and encouragement to return the test. If patients still do not return the FIT test, existing national and local guidelines should be used to assess risk of CRC. A limited evidence base suggested that people from ethnic minorities may be less likely to return the test possibly due to hygiene concerns. Clinical advisors to the EAG noted the use in primary care of software (e.g., AccuRx) to send text message reminders and list non-completers for follow up.

1.2.4 DAP50 proposed use of FIT

In this assessment, FIT will be evaluated as an adjunct to clinical assessment to guide referral of a symptomatic population to the suspected CRC pathway. Consistent with the ACPGBI/BSG guideline, this population includes both those meeting NG12 criteria for an urgent 2WW suspected CRC referral, and those meeting DG30 criteria for a FIT test. Patients would receive the test in primary care, and the result of the test would be used to determine who would proceed to secondary care and who would be followed up in primary care with safety netting. However, FIT results should not be the sole arbiter of referral (i.e., persons with a negative FIT could still be referred directly to urgent 2WW suspected CRC pathway based on ongoing clinical concerns), and this will be explored with our clinical advisors.

1.3 Clear definition of the intervention

1.3.1 FIT testing

FIT is available as quantitative tests (using immunoturbidimetric or enzyme-linked immunosorbent assay [ELISA] methods to measure haemoglobin concentration) or qualitative tests (using immunochromatographic test devices to detect haemoglobin). In line with DG30 this evaluation will focus on quantitative FIT.

Immunoturbidimetric FIT contains particles which are coated in antibodies specific to human haemoglobin. The antibodies bind to haemoglobin present in the faecal sample creating complexes which are detected using turbidimetry.

ELISA FIT uses antibodies specific to human haemoglobin to bind haemoglobin in the faecal sample to the surface of microtiter wells. This is then treated with chemicals to produce a colour change. The intensity of the colour is proportional to the amount of haemoglobin in the sample. Some assays may also include antibodies for human haptoglobin. Haptoglobin is a protein produced by the liver which binds to haemoglobin making it less likely to break down during transit through the gastrointestinal tract. The detection of haptoglobin is claimed to increase the likelihood of detecting lesions higher in the colon.

Different FIT may report outcomes using either the concentration of haemoglobin in the sampling device buffer (nanograms Hb/mL buffer) or as concentration of haemoglobin by mass of faeces (μg Hb/g). As the amount and type of buffer used varies between manufacturers, the World Endoscopy Organization's expert working group on FIT for CRC screening recommended that μg Hb/g should be used as a standard measure that can be compared easily between tests.[12]

1.3.2 Strategies and thresholds for using FIT as a triage tool

Since the test is quantitative, thresholds may be varied to achieve optimal clinical and cost effectiveness outcomes with respect to COL capacity, quality-adjusted life years (QALYs) or Net Health Benefit (NHB). The testing strategies (i.e., threshold used) to be assessed will be determined by the available literature.

Strategies using one FIT threshold will be investigated, where FIT above a threshold will result in referral to the urgent 2WW suspected CRC pathway, whilst FIT below the threshold will result in safety netting (see Section 1.9.5) which may include referral to other pathways such as the non-specific cancer pathway, urgent non-cancer referrals or the 18WW non-urgent lower GI pathway. Strategies using two FIT thresholds (t_1 and t_2) to determine management pathways will also be considered (see Figure 1). In these strategies, people with FIT over t_1 would be referred to the urgent 2WW suspected CRC pathway and people with FIT below t_2 would be followed-up with safety netting in primary care. People with FIT between t_1 and t_2 may receive either a non-urgent referral or other types of follow-up (see Section 1.9.1).

There are several tests within the scope of this assessment. These are described in sections 1.3.3 to 1.3.9.

1.3.3 HM-JACKarc system

The HM-JACKarc system (Hitachi Chemical Diagnostic Systems Ltd, Alpha Laboratories) is a fully automated quantitative immunoturbidimetric FIT system. The system comprises a sample collection device (designed to measure 2 mg of faeces) which contains 2 mL of stabilising buffer, latex agglutination reagent, and buffer solution. The assay is compatible with the HM JACKarc analyser, which can process up to 200 samples per hour, with a maximum capacity of 80 samples per run.

1.3.4 FOB Gold

FOB Gold (Sentinel/Sysmex) is an automated quantitative immunoturbidimetric FIT system. It comprises faecal sample collection tubes (the SENTiFIT pierceTube faecal collection device) which collect 10 mg of faeces in 1.7 mL of buffer, and latex agglutination reagent. The FOB Gold kit is compatible with Sentinel's own SENTiFIT analyser as well as those manufactured by 5 other companies. The performance characteristics of the assay vary depending on which analyser is used. The throughput of the test is dependent upon the clinical chemistry analyser used to process the samples, but 270 samples can be run per hour on the SENTiFIT 270.

1.3.5 OC-Sensor

The OC-Sensor (Eiken Chemical/MAST Diagnostics) is a quantitative immunoturbidimetric FIT. It comprises faecal sample collection tubes, latex reagent, and buffer. The OCAuto sampling bottles can hold 10 mg of faeces.

The test can be run on either the OC-Sensor PLEDIA or the OC-Sensor iO analyser, which differ in the number of samples they are able to process. The OC-Sensor PLEDIA can process up to 320 samples per hour, with a capacity of 200 samples per run. The OC-Sensor iO can process up to 88 samples per hour with a maximum capacity of 20 samples per run.

[redacted]

1.3.6 NS-Prime

The NS-Prime (Alfresa/Abbott) is an automated quantitative immunoturbidimetric FIT system. The NS-Prime comprises a specimen collection container which collects 10 mg of faeces in 1.9 mL of buffer solution (Carroll et al. 2014). The test is run on the NS-Prime clinical chemistry analyser.

The NS-Prime haemoglobin reagent is specific to the NS-Prime analyser and cannot be used on other platforms. The NS-Prime analyser can run up to 220 samples at the same time, processing 300 tests per hour.

1.3.7 IDK TurbiFIT

The IDK TurbiFIT assay (Immundiagnostik) is an immunoturbidimetric FIT compatible with a range of automated clinical chemistry analysers from 16 manufacturers. The TurbiFIT kit comprises reagents, control samples, and calibration samples. IDK TurbiTUBE sample collection devices are available separately, which collect 15 mg of faeces in 1.5 mL of buffer. The performance characteristics and throughput of the assay vary depending on which analyser is used.

1.3.8 IDK Hemoglobin (human) and hemoglobin/haptoglobin complex ELISA tests

The IDK hemoglobin (human) ELISA (Immundiagnostik) is an immunoassay for the quantitative determination of human haemoglobin in faeces. It consists of:

- a microtiter plate, pre-coated in antibodies
- buffers for washing, extraction, and sample dilution
- conjugate peroxidase-labelled antibodies
- standards and controls
- tetramethylbenzidine substrate (to induce the colour change)

The test requires an ELISA plate reader with a photometer (Dynex DS2 and DSX systems) to determine the result. The throughput of the test is dependent upon the clinical chemistry analyser used to process the samples.

The company also produces the IDK hemoglobin/haptoglobin complex ELISA, which is similar but uses anti-haptoglobin antibodies in the coated microtiter plate. The company recommends using this test in addition to a haemoglobin test to improve sensitivity for detection of bleeding adenomas or cancers of the upper intestine.

1.3.9 QuikRead go iFOBT

The QuikRead go (Aidian) is a point-of-care analyser that can be used for a number of different diagnostic tests, including the immunochemical faecal occult blood test (iFOBT) which is an immunoturbidimetric test. The kits contain reagent capsules and buffer in prefilled cuvettes. Faecal sampling sets and control materials are supplied separately. A single sample can be run at a time, and the test takes less than 2 minutes for the result to be displayed.

Table 1: Summary of interventions (adapted from Table 1 from the NICE scope[13])

Test (see sections 2.2.1 to 2.2.7 for more detail)	**Test principle**	**Analyser compatibility**	**Sample size required** (mg)	**Measuring range** (µg Hb/g)	**Limit of detection** (µg Hb/g)	**Limit of quantitation** (µg Hb/g)	**Throughput**
HM-JACKarc	Immunoturbidimetry	HM JACKarc analyser	2	7 to 400	0.6	1.25	200 samples per hour
FOB Gold	Immunoturbidimetry	Various	10	Varies according to the analyser used	Varies according to the analyser used	Varies according to the analyser used	Dependent on the analyser used
OC-Sensor	Immunoturbidimetry	OC-Sensor PLEDIA	10	2 to 50,000	2	2	320 samples per hour
	Immunoturbidimetry	OC-Sensor iO	10	2 to 200	2	4	88 samples per hour
NS Prime	Immunoturbidimetry	NS-Prime analyser	10	4 to 240	4	10	300 tests per hour
IDK TurbiFIT	Immunoturbidimetry	Various	15	Varies according to the analyser used	Varies according to the analyser used	Varies according to the analyser used	Dependent on the analyser used
IDK	ELISA	Various	15	0.18 to 50	0.15	0.18	Dependent on

Hemoglobin ELISA		(ELISA plate reader with a photometer (Dynex DS2 and DSX systems))					the analyser used
IDK Hb/Hp complex ELISA	ELISA		15	0.25 to 50 µg HbHp/g	0.16 µg HbHp/g	0.25 µg HbHp/g	Dependent on the analyser used
QuikRead go iFOBT (point-of-care test)	Immunoturbidimetry	QuikRead Go analyser	10	10 to 200	2.5	9.5	Less than 2 minutes per test.

Information provided by companies to NICE or taken from the test's instructions for use document or website. ELISA, enzyme-linked immunosorbent assay; Hb, haemoglobin; Hp, haptoglobin. Accuracy should be analysed according to analyser used if data is available.

1.4 Populations and relevant subgroups

The population is people presenting to primary care with symptoms or signs indicating a risk of CRC, as defined by NG12 and DG30.

Certain symptoms may indicate patients should be referred directly to the urgent 2WW suspected CRC pathway (people with palpable rectal or anal mass or anal ulceration) and these patients are excluded from the scope. However, see Section 1.9.3, where their potential inclusion in the modelling is described.

Faecal haemoglobin levels are thought to differ according to certain patient characteristics. Different cut-off values may be needed according to the following characteristics:

- Age
- Sex
- Ethnicity
- People taking medications or with conditions which increase the risk of gastrointestinal bleeding
- People with blood disorders (e.g., beta thalassemia) that could affect the performance of the test
- People with anaemia (including iron deficiency anaemia)

Although FIT is proposed to be offered to the population with CRC symptoms, it is possible that introduction of the test would have an indirect impact on people waiting for non-urgent referral to gastroenterology services and/or COL for non-cancer conditions.

1.5 Relevant comparators

Current practice corresponds to standard care according to current NICE guidelines NG12 and DG30 (see section 1.2.1). This includes:

- Clinical assessment and referral for further investigation in secondary care
- Use of FIT (threshold of 10 µg Hb/g) to guide referral only for those with 'low-

risk' symptoms without rectal bleeding (in line with NICE guideline DG30)

Feedback from clinical experts and stakeholders is that stratification by symptoms is a poor predictor of risk of CRC. Any resulting guidance that differentiates between the risk groups currently defined in NICE guidance would not address this problem. Therefore, despite the possibility of differential cost-effectiveness by subgroup, the intervention arm will not subgroup according to NG12 high-risk and DG30 low-risk categories and will not exclude those with active rectal bleeding, to avoid making recommendations according to symptom-based criteria. Consequently, the comparator is a blended group of people who would currently be considered under the guidance of NG12 and DG30.

The aim of the modelling is to determine the most cost-effective FIT strategy to reduce the number of people without significant bowel pathology who are referred to the suspected CRC (2WW) pathway, taking into consideration potential COL capacity constraints for urgent and non-urgent referrals. More details about the comparators that may be used in the modelling are given in Section 3.3.2.2.

1.6 Reference standard

For diagnostic test accuracy studies, the ideal reference standard is full colonic imaging using COL or CTC. Other reference standards (e.g., long-term follow-up; differential reference standards based on FIT result) will be considered where data using the preferred reference standard is unavailable and bearing in mind the potential limitations of these (e.g., long-term follow-up may detect CRC that was not present at the time of the index test).

1.7 Healthcare setting

The project relates to the use of FIT in primary care.

1.8 Outcomes

There are unlikely to be any end-to-end studies that compare FIT to current practice in the specific population of interest.

Intermediate outcomes of interest may include:

- Diagnostic accuracy at different FIT thresholds (NB we will only include studies reporting CRC diagnostic test accuracy, but will also extract data relating to IBD and AAs from these studies where it is reported)
- Risk of CRC (and IBD and AAs) in relevant subgroups according to FIT threshold
- Test failure rates
- Prognostic implications of false-negative results
- Uptake (completion) of FIT in primary care
- Number/proportion of people referred to secondary care
- Number/proportion of people followed up in primary care
- Duration of validity of negative test (implications for follow-up)
- Number/proportion of urgent (2WW suspected cancer) specialist appointments
- Number/proportion of urgent (2WW suspected cancer) COL/CTCs
- Number/proportion of non-urgent COL/CTCs
- Time to COL/CTC
- Time to diagnosis of CRC or other conditions
- Number/proportion of COL/CTCs that do not detect CRC
- Number/proportion of COL/CTCs that do not detect significant bowel pathology
- Number/proportion of people presenting to emergency departments with symptoms of CRC

Clinical outcomes for consideration may include:

- Number of CRC diagnoses
- Number/proportion of CRC diagnoses from urgent referrals
- Stage of detected cancers
- Number/proportion of people identified with other bowel pathologies
- Number/proportion of people with advanced adenomas detected, or detected and treated
- Morbidity including adverse events associated with COL
- Mortality

Patient-reported outcomes for consideration may include:

- Health-related quality of life
- Anxiety associated with waiting for referral or test results due to diagnostic delays, and further diagnostic workup
- Preference for FIT versus COL

Costs will be considered from an NHS and Personal Social Services (PSS) perspective. Costs for consideration may include:

- Cost of equipment, reagents, and consumables for FIT
- Cost of staff and associated training
- Medical costs arising from testing and care including further follow-up and safety netting
- Medical costs arising from adverse events which arise from testing or further diagnostic work up, including those associated with false test results and inappropriate treatment.

The cost-effectiveness of FIT versus usual practice will be expressed in terms of the incremental cost per QALY gained. NHB will be used when comparing multiple interventions. A lifetime horizon will be used. Further details of the proposed health economic analysis are presented in Section 3.

1.9 Other considerations

There is known to be heterogeneity within care pathways across the country. Patient care pathways on which the modelling will focus should be identified via consultation with clinical advisors. We will obtain expert opinion to understand (1) what should happen (according to guidelines and clinical opinion), (2) what the current heterogeneity is in care pathways and (3) what happens in the majority of places. This information will be used to determine what pathways to model for the base case and if any additional pathways should be modelled in scenario analyses if data and time allows.

Amongst the “other considerations” discussed in Sections 1.9.1 – 1.9.9, modelling

scenario analyses are planned for items 1.9.1, 1.9.1.1, 1.9.2, 1.9.4, 1.9.5 and 1.9.6.

1.9.1 FIT threshold for referral

The FIT cut-off recommended in DG30 was 10µg/g, as the committee concluded this gave the test enough sensitivity to reliably rule out CRC in the low-risk population. FIT thresholds may be varied for two reasons:

- To optimise the treatment pathway for clinical effectiveness (QALYs) or cost-effectiveness (in terms of NHB) and to investigate impact on numbers/proportions of referrals
- Because faecal haemoglobin levels are thought to differ according to certain patient characteristics. Different cut-off values may be needed for these subgroups to avoid potential equity issues:
  - Age groups (to be determined by available evidence)
  - Sex (male/female)
  - Ethnicity groups (to be determined by available evidence)
  - People taking medications or with conditions which increase the risk of gastrointestinal bleeding
  - People with blood disorders (e.g., beta thalassemia) that could affect the performance of the test
  - People with anaemia (including iron deficiency anaemia)

Both reasons for threshold alteration will be considered in the assessment.

1.9.1.1 Use of two FIT thresholds to guide referral, and the intermediate group pathway

As described in Section 1.3 and Figure 1, two FIT thresholds could be used to define low, intermediate, and high-risk populations. In this strategy, people in the intermediate risk group may have more intensive monitoring of their condition than in the low-risk group or be referred to a specialist safety netting pathway (see Section 1.9.5).

The impact of using different thresholds, either to define two or three risk groups, will be investigated in the assessment. Currently, the management pathway for the

intermediate group is unclear. The impact of different assumptions about this will be explored within scenario analyses (see Section 3.3.2) and may include some of the strategies described in the section on safety netting (Section 1.9.5). Assumptions will be based on expert advice and also informed by pathways used during the peak of the COVID pandemic.

1.9.2 Measurements and diagnostic test accuracy of different tests and analysers

Different tests, different analysers and different combinations of tests and analysers (see Table 1 and Table 2) may have different measuring ranges, may give different absolute measurements, and may have different test accuracy. This will be addressed by primarily focussing on evidence relating to the specific tests and analysers defined in the scope, and by considering each test or combination of test and analyser individually (see Section 2.10). The generalisability of data between tests may be considered in modelling scenarios (see Section 2.2) and may be informed by equivalence data submitted by companies (see Section 1.10).

1.9.3 Use of FIT alongside bypass referral

Patients who have bypass symptoms are not part of the decision problem population. Clinical experts also advised NICE that (compared to 2019, when this topic was first scoped) rectal bleeding would no longer be considered a reason to bypass FIT. However, presence of a palpable rectal or anal mass, or anal ulceration may be reason to move straight to a 2WW referral bypassing FIT, and this is referred to in both the ACPGBI/BSG guideline and the NHS England letter. Some clinical experts said that FIT could still be useful alongside referral, to help choose the method of further investigation, and may be required by some secondary care centres. Although the bypass symptoms are not part of the decision problem population, we may include them to inform the capacity that is available for COL. The appraisal team will consult experts to determine what current care should be. This may be considered within the modelling to inform capacity estimates.

1.9.4 Dual testing

Dual FIT is using two samples from different bowel movements rather than a single

sample from one bowel movement. Based on clinical expert opinion, people would be referred to the suspected cancer pathway if either FIT sample were positive. Dual FIT may result in fewer false negative results, more false positive results, higher costs of FIT testing, more need for laboratory processing of FIT tests and a change in FIT compliance. Note that this is a different scenario to using FIT as part of a safety netting programme (see section 1.9.5, "offering second FIT"). Where data allows, the impact of dual testing will be investigated in the assessment.

1.9.5 Safety netting

Safety netting refers to processes used in the diagnostic pathway to avoid missing disease (cancer or otherwise) in patients with a negative FIT result. This pathway appears to be highly heterogeneous and is currently evolving.

The assessment for DG30 assumed that patients who had a false negative FIT result would re-present in primary care within 12 months with ongoing or worsening symptoms, with a delay to diagnosis. DG30 modelling assumed the following for safety netting (persons with 'negative FIT'): (i) if they had cancer, they would have a delay in diagnosis of less than 12 months, (ii) for those without cancer a proportion would have persistent symptoms: some would receive COL and some would receive a repeat FIT. For (ii) proportions were estimated based on clinical opinion (two clinicians who provided quite different estimates, Table 26 of DG30[14]); the EAG assumed 32% went on to receive COL and 20% had repeat FIT.

NG12 recommends safety netting for people with symptoms associated with an increased risk of cancer who do not meet the criteria for referral or other investigative action. This may be planned within a timeframe agreed with the person, or initiated by the person if their symptoms recur, persist, or worsen. The guideline also highlights the possibility of false negative results from FIT. The ACPGBI/BSG guidance recommends that safety netting protocols should include advice and strategies for the diagnosis of colorectal and extracolonic cancers, as well as other serious gastrointestinal conditions.

The recent NHS England letter stated that the ACPGBI/BSG guideline should be implemented in full and provided recommendations for safety netting. These stated that clinical teams should consider:

- "*Providing the patient with clear information about who to contact if they develop new symptoms or if their existing symptoms worsen.*
- *Using advice and guidance via eRS to guide management of patients with persistent or troublesome symptoms.*
- *Offering a second FIT if ongoing clinical concerns remain.*
- *Referral to a non-specific-symptoms urgent cancer pathway, if appropriate and there are ongoing concerns about possible cancer.*
- *Management of FIT negative patients in an outpatient setting following referral on a non-urgent pathway. For example, the North Central London Cancer Alliance has developed a FIT negative, non–urgent referral pathway, as has Oxford University Hospitals NHS Foundation Trust.*"

The eRS system is used in some areas as a means of communication between primary and secondary care for advice and guidance, whilst in others it may only be used to make and track referrals. Other methods of communication may be used between primary and secondary care for advice and guidance.

Safety netting will be included as part of the diagnostic pathway of patients with negative FIT results in this assessment. In this assessment, we include referral to the 18WW pathway for patients who do not meet criteria for the 2WW pathway (either NG12 or FIT threshold) as part of safety netting. Guidelines all leave some room for interpretation and practice may be heterogeneous as a result. We will explore safety netting options with our clinical advisors. Where data allow, scenario analyses could investigate different assumptions relating to the intensity of safety netting, and may incorporate, for example, two extremes: a low intensity assumption as in DG30; and a high intensity assumption, based on the NHS England letter, to give the committee a range of cost effectiveness estimates and numbers of additional referrals to consider (see Section 3.3.2 and 3.3.6).

It should be noted that it is unclear at this stage whether all inputs required for the modelling of scenarios will be available from published or unpublished sources. Studies reporting on different safety netting protocols will be identified and flagged for use in the economic model but will not be the subject of full systematic review and synthesis. Attempts will be made to obtain data directly from key research groups (e.g., North Central London Alliance, Oxford University NHS Foundation Trust), but if all required data are not available, or are unavailable in time for incorporation in the assessment, assumptions may have to be made about diagnostic accuracy, delays in time to diagnosis and costs relating to these strategies. This may introduce uncertainty to the cost-effectiveness estimates. Equally, low quality evidence (e.g., small sample sizes; incomplete data collection) may introduce uncertainty.

1.9.6 Other conditions with gastrointestinal symptoms

Patients presenting with symptoms of CRC may have other gastrointestinal pathologies such as IBD (CD or UC), diverticular disease or AAs. COL is required to diagnose IBD, and to identify and treat AAs. A patient with any of these conditions will only receive a diagnosis and treatment if they receive a lower gastrointestinal (GI) referral (either an urgent 2WW referral for suspected CRC, a 18WW referral or another type of lower GI referral). Note that for these patients the pathways '18WW referral' and 'safety netting' may result in a delay in diagnosis. A delay in diagnosis for IBD may worsen QoL and patient outcomes so IBD has been included within the scope of the modelling. AAs are largely asymptomatic, but some may eventually develop into CRC if not treated. A finding of AA during COL is largely incidental (since adenoma symptoms did not trigger the investigation, as they are largely asymptomatic), but a benefit of receiving COL. If data and time allows, AAs will be included within the scope of the modelling in the base case but excluded within a scenario analysis.

The COLOFIT project conceptual modelling has opted to explicitly include IBD (both CD and UC) within the model due to the known impact of a delayed diagnosis on prognosis, costs and quality of life. Other bowel diseases were not modelled explicitly due to a lack of clarity around whether a diagnostic delay is likely to cause harm. It is

proposed that a similar approach be taken for DAP50. This will require estimates of (1) FIT sensitivity to IBD, and (2) undiagnosed IBD prevalence in the population. These data will primarily be sought from diagnostic test accuracy studies of FIT for CRC but may be supplemented with additional focussed searches if required and expert advice.

1.9.7 Urgent 2WW suspected CRC pathway and secondary care management

We will aim to understand heterogeneity in current practice regarding what happens upon a referral to the urgent 2WW suspected CRC pathway and represent costs and capacity relating to this in the modelling (see Section 3.3.2).

1.9.8 Non-urgent referral pathway

It is currently unclear what the non-urgent referral pathway entails. We will aim to understand what happens upon a referral to the non-urgent pathway, and which patients are referred to this pathway, and will represent costs and capacity relating to this in the modelling.

1.9.9 Non-completers of FIT tests

A proportion of patients will not return their FIT tests. DG30 reported that this ranged from 41% to 98% for OC Sensor, 56%-66% for HM-JAKarc. We will aim to understand heterogeneity in current practice regarding what happens to these patients and represent costs and capacity relating to this in the modelling. Currently, we have heard from clinical advisors that eRS may be used to send automatic text reminders to patients, and that some patients may get a direct referral to the urgent 2WW CRC pathway.

1.10 Areas that are outside the scope of the appraisal and therefore do not require any detailed assessment (e.g., key factors for which evidence is already accepted).

Evidence on equivalence of devices will not be sought or statistically synthesised by the EAG. If evidence is submitted by companies relating to equivalence, it may be assessed and used to inform modelling scenarios (see Section 1.9.2).

Development of a risk prediction model using FIT and clinical characteristics is not within the scope of the assessment. This type of work is being conducted by other

groups (e.g., NICE FIT group, COLOFIT). A review of risk prediction models is also not within the scope of this assessment, since this work is being conducted by the COLOFIT group.

Colon capsule endoscopy is an emerging technology that may be in use in some areas. Due to a lack of evidence in the symptomatic population, and since it is not widely used, it will not be considered in this assessment.

2. Report methods for assessing the outcomes arising from the use of the interventions

A systematic review will be conducted to identify clinical efficacy and diagnostic test accuracy studies of relevance to the decision problem. Clinical efficacy studies refer to "end-to-end" studies which compare two different testing strategies using a randomised control trial (RCT) design, whereas diagnostic test accuracy studies refer to studies that report intermediate outcomes such as sensitivity and specificity, using a cohort or cross-sectional design.

Summary of our approach: The ACPGBI/BSG guidance was based on a recent systematic review of the literature relating to clinical efficacy and diagnostic test accuracy. Some of the authors of this review are clinical advisors to the EAG and have committed to sharing their review work with the EAG for the purposes of completing DAP50. As such, DAP50 will comprise an update of the ACPGBI/BSG review, where the reviews overlap. We will also consult the EAG report for DG30 for studies. Where the scope of the ACPGBI/BSG review is narrower than that of DAP50, additional review work will be conducted to fill these gaps. For example, where data were not extracted for all thresholds reported in a study, the original study will be revisited to perform *de novo* data extraction. All data extractions and quality assessments submitted by ACPGBI/BSG group will be checked by ScHARR and any discrepancies resolved through discussion with the ACPGBI/BSG authors.

A review protocol will be prepared and prospectively submitted to PROSPERO (https://www.crd.york.ac.uk/prospero/).

2.1 Population

Studies will be included that recruited people presenting to primary care with signs or symptoms indicating a risk of CRC. Signs and symptoms of CRC are defined as those described in NG12 and DG30. Studies will be included if they recruited patients either with NG12 high-risk symptoms, or with DG30 low-risk symptoms or with both. Ideally studies would exclude patients who would bypass FIT testing due to the presence of a rectal or anal mass, or anal ulceration. Studies reporting data relating to the subgroups

specified in the population section (Section 1.4) of the decision problem (e.g., age, sex) will also be included.

We anticipate that we will encounter the following patient recruitment strategies amongst the literature:

1. Recruitment of patients in primary care: either a) all patients presenting to primary care with signs and symptoms of CRC regardless of low-risk or high-risk status or b) patients exclusively with, or subgrouped by, low-risk (DG30) or high-risk (NG12) symptoms are included. Recruitment of group a) is the ideal recruitment strategy.
    - Ideally all patients would be followed up with full colonic imaging (COL/CTC), but it is possible that these studies will have a differential reference standard based on the results of the index test. For example, FIT positive patients get full colonic imaging (COL/CTC), whilst FIT negative patients get safety netting and long-term follow up. Not following all patients up with full colonic imaging (COL/CTC) as the reference standard but using long term follow-up instead may lead to; a) missed false negatives, i.e. patients who had CRC but were not picked up by FIT are not identified by the reference standard, or b) incorrect FIT false negatives, i.e. patients who did not have CRC at the time of the FIT index test and were true negative, but where CRC has emerged during a long follow-up. It is thought that most patients who had a false negative FIT test will re-present within 3 months. Without knowing FN, TN cannot be known either. For these types of studies, we will consider the length of follow up and the likely impact on results (see Section 2.3).
2. Recruitment of patients in secondary care with high-risk (NG12) symptoms. This may miss some patients referred in primary care to other diagnostic pathways, and therefore may miss a small number of false negative and true negative patients but is likely to only have a small impact on estimates of test accuracy.
    - It is likely the reference standard will be COL/CTC (or other imaging modality if indicated) for all secondary care patients.

3. Recruitment of patients in secondary care who have been referred from primary care, regardless of low-risk (DG30) or high-risk (NG12) symptoms.
    - This recruitment strategy is potentially problematic since patients presenting to primary care who are not referred are not included in the study. This may mean that DG30 patients (or any patients not referred if the study did not use DG30/NG12 guidelines) negative by FIT have been excluded, and an estimate of false negatives and true negatives would be absent.

We will consider the impact of the recruitment strategy and reference standard used on the relevance of the data to the assessment. Where studies recruit populations that differ from the desired population in any respect, they may be included where no data is identified in the desired population and where generalisability is thought to be reasonable. We will base this decision on the proportion of out-of-scope participants, and on statistical consideration and clinical opinion as to the likely impact of their inclusion. This may be especially relevant when seeking information relating to subgroups (e.g., patients with characteristics such as age or sex that may require a different threshold), sensitivity (e.g., the impact of an imperfect reference standard; the impact of the specific test on diagnostic test accuracy) or for scenario analyses.

2.2 Interventions

Studies will be included if they report data using any of the test-analyser combinations listed in Section 1.3, and in Table 1. These are:

- HM JACKarc with HM JACKarc analyser, manufactured by Hitachi Chemical Diagnostic Systems Ltd and distributed by Alpha Laboratories,
- FOB Gold manufactured by Sentinel and distributed by Sysmex, for use with a range of clinical chemistry analysers, including those supplied by Siemens, Beckman Coulter, and Abbott
- OC-Sensor with analysers PLEDIA, iO or [redacted], manufactured by Eiken Chemical and distributed by MAST Diagnostics
- NS Prime with NS-Prime clinical chemistry analyser, manufactured by Abbott and distributed by Alfresa

- IDK TurbiFIT for use with various analysers, manufactured and distributed by Immundiagnostik
- IDK Hemoglobin ELISA for use with ELISA plate readers with a photometer (Dynex DS2 and DSX systems), manufactured and distributed by Immundiagnostik
- IDK Hb/Hp complex ELISA for use with ELISA plate readers with a photometer (Dynex DS2 and DSX systems), manufactured and distributed by Immundiagnostik
- QuikRead go iFOBT (point-of-care test) for use with the QuickRead Go analyser, manufactured, and distributed by Aidian.

Each test and test-analyser will be considered individually. Where data is not available on a specific test, evidence of equivalence will be considered if submitted by the company and evidence from other tests may be used to inform scenario analyses in the model. For example, where data relating to different thresholds by age is only available for a limited number of analysers, a sensitivity analysis could be undertaken within the model in which it is assumed that the relative impact of a specific subgroup can be generalised for all analysers within the modelling. Outputs from the statistical analysis may help to inform assumptions of equivalence between tests. Since the IDK Hb/Hp complex ELISA test measures Hp/Hb, not Hb like the other tests, thresholds and evidence are likely to be unique to this test.

Data relating to all thresholds will be sought, to facilitate an evidence-based selection of thresholds for inclusion in the statistical analysis and economic modelling.

Studies reporting dual testing using any of these analysers will be included in the review.

2.3 Comparators

For the review of clinical efficacy: In the first instance, the review will seek end-to-end studies that compare one diagnostic strategy to another. Studies will be eligible for inclusion if they compare FIT to standard care. Standard care currently includes clinical

assessment in primary care and referral to secondary care for imaging (COL or CTC) based on gastrointestinal symptoms alone (NG12 high-risk patients) and referral based on gastrointestinal symptoms and FIT result (DG30 low-risk patients). It is considered unlikely that there will be any end-to-end clinical efficacy studies.

For the review of diagnostic test accuracy studies and comparative diagnostic test accuracy studies: The ideal reference standard is full colonic imaging with COL or CTC. In the first instance, studies where this reference standard has been applied to all patients will be sought. Other reference standards (e.g., long-term follow-up; differential reference standards based on FIT result) will be considered where data using the preferred reference standard is unavailable and bearing in mind the potential limitations of these (e.g., long-term follow-up may detect CRC that was not present at the time of the index test; most FIT false negative patients re-present within 3 months). It may be necessary to subgroup studies according to the reference standard used. Categories will be based on the available literature but may be in accordance with the categories used in the ACPGBI/BSG guideline (tier 1, >90% underwent full colonic imaging with COL or CTC; tier 2, <90% underwent full colonic imaging or patients were followed up for more than 3 months). It is thought unlikely that studies using colon capsule colonoscopy as the reference standard will be found, but if they are we will consider their inclusion based on clinical advice regarding whether colon capsule colonography is a good reference standard (i.e., has diagnostic accuracy comparable to COL or CTC).

Comparative diagnostic test accuracy studies are those that compare two or more of the tests or test-analyser combinations listed in Section 1.3 to each other. Where only one test within the scope of this assessment is reported, the study would be treated as a diagnostic test accuracy study.

2.4 Outcomes

For the review of end-to-end clinical efficacy studies, the following outcomes will be eligible for inclusion:

- Number of CRC diagnoses

- Number/proportion of CRC diagnoses from urgent referrals
- Stage of detected cancers
- Number/proportion of people identified with other bowel pathologies
- Number/proportion of people with advanced adenomas detected, or detected and treated
- Morbidity including adverse events associated with COL
- Mortality
- Health-related quality of life
- Anxiety associated with waiting for referral or test results due to diagnostic delays, and further diagnostic workup
- Preference for FIT versus COL
- Risk of CRC (and IBD and AAs) in relevant subgroups according to FIT threshold
- Test failure rates
- Prognostic implications of false-negative results
- Uptake (completion) of FIT in primary care, to include with respect to cultural, demographic, or socioeconomic factors
- Number/proportion of people referred to secondary care
- Number/proportion of people followed up in primary care
- Duration of validity of negative test (implications for follow-up)
- Number/proportion of urgent (2WW suspected CRC) specialist appointments
- Number/proportion of urgent (2WW suspected CRC) COL/CTCs
- Number/proportion of non-urgent COL/CTCs
- Time to COL/CTC
- Time to diagnosis of CRC or other conditions
- Number/proportion of COL/CTCs that do not detect CRC
- Number/proportion of COL/CTCs that do not detect significant bowel pathology
- Number/proportion of people presenting to emergency departments with symptoms of CRC

In the likely event that there are no end-to-end clinical efficacy studies, these outcomes may be either reviewed as model parameters, reviewed via diagnostic test accuracy studies, or be outputs from the model.

In the review of diagnostic test accuracy studies, the following outcomes will be eligible for inclusion and extraction:

- Number of true-positives, true-negatives, false-positives, and false-negatives, only where all four statistics are reported or can be calculated (these data allow the calculation of sensitivity, specificity, positive and negative predictive values if these are required), for CRC, IBD and AAs. NB only studies recruiting patients with signs and symptoms of CRC will be included, and IBD and AA data will only be extracted from studies that used FIT in the context of identifying CRC. It is anticipated that CRC, IBD and AA data will be reported in a single publication, but where this is split across publications for a given study, data will be extracted from separate publications.
- Receiver-operating-characteristic (ROC) curves for digitisation and use in the synthesis, if required
- Other outcomes as listed for the clinical efficacy studies will be identified if these are likely to be useful for informing the parameters of the health economic model. For example:
    - risk (prevalence) of CRC, IBD and AA for the whole population and relevant subgroups
    - test failure rates
    - uptake of FIT in primary care (non-completion of tests)

2.5 Study design

For the review of end-to-end clinical efficacy studies, RCTs or controlled clinical trials (CCTs) will be eligible for inclusion.

For the review of diagnostic test accuracy and comparative diagnostic test accuracy, only cohort or cross-sectional studies that recruited patients regardless of eventual diagnosis will be included (i.e., studies that avoided a case-control design).

Relevant systematic reviews identified during study selection will be used to check for additional studies and may be used for data extraction (see Section 2.8).

Studies not published in English language will be included if sufficient data can be extracted from non-English language full-texts, or from an existing English language abstract. Conference abstracts and non-peer-reviewed reports will only be included if the data are presented in a succinct and accessible manner (e.g., a manuscript prepared for submission to a journal), if sufficient methodological details are reported to allow critical appraisal of the study quality, and if results are reported in sufficient detail. Where there are gaps in the available literature, exclusion criteria for conference abstracts and non-English language papers may be relaxed.

2.6 Search strategy

Searches will be conducted to identify evidence on the intervention (FIT assays) and target condition (CRC), as recommended in the Centre for Reviews and Dissemination (CRD) guidance for undertaking reviews in health care[15] and the Cochrane Handbook for Diagnostic Test Accuracy Reviews.[16]

Searches will be based upon and update those conducted for the ACPGBI/BSG review (31/3/22), which was in turn based upon the searches for DG30 (March 2016). Additional focused searches will be conducted should any discrepancies between the scope of DAP50 and the ACPGBI/BSG review be identified.

Searches will consider generic and product names for the intervention. An example search strategy is supplied in Appendix 1. Search strategies will be optimised for each database, to include:

- MEDLINE (Ovid) including Epub ahead of print, In-Process Citations and Daily Update (Ovid)
- EMBASE (Ovid)
- Cochrane Database of Systematic Reviews (CDSR) (Wiley)
- Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley)
- International Network of Agencies for Health Technology Assessment (INAHTA) Publications (Internet) http://www.inahta.org/publications/

- NIHR Health Technology Assessment Programme (Internet)
- Aggressive Research Intelligence Facility (ARIF) database (Internet) http://www.birmingham.ac.uk/research/activity/mds/projects/HaPS/PHEB/ARIF/ind ex.aspx
- PROSPERO (International Prospective Register of Systematic Reviews) (Internet) http://www.crd.york.ac.uk/prospero/
- Completed and ongoing trials will be identified by searches of the following resources:
- NIH ClinicalTrials.gov (http://www.clinicaltrials.gov/)
- EU Clinical Trials Register (https://www.clinicaltrialsregister.eu/ctr-search/search)
- WHO International Clinical Trials Registry Platform (ICTRP) (http://www.who.int/ictrp/en/)

Key conference proceedings, to be identified in consultation with clinical experts, will be screened for the last three years. No language restrictions will be applied.

Identified references will be downloaded into reference management software for further assessment and handling. Reference lists in included articles and relevant systematic reviews will be checked for additional studies. The final list of included papers will also be checked on PubMed for retractions, errata, and related citations. Clinical experts will be consulted for any missed studies.

If differences between the scope of the DG30 or ACPGBI/BSG reviews and our own review are identified, the list of studies excluded at full text in the DG30 and ACPGBI/BSG reviews will be interrogated to ensure that no studies of relevance to this review have been excluded. Studies included in either review will be checked for subgroup and threshold data relevant to our review that were not relevant to their review. Submissions from companies will also be checked for additional data.

2.7 Study selection

Studies will be selected for inclusion in the review if they meet the inclusion criteria detailed in Sections 2.1 to 2.5. Titles and abstracts will be considered for inclusion against the criteria by one reviewer, with a 10% sample checked by a second reviewer. This will be conducted on the first 10%, and before the remainder are screened, and will be repeated in 5% increments until a high level of agreement regarding included studies is achieved. Full texts will be obtained and considered for inclusion by one reviewer, with decisions checked by a second reviewer. Any discrepancies will be resolved through discussion, or with reference to a third reviewer (e.g., clinical advisor).

2.8 Data extraction strategy

The data extraction form used by ACPGBI/BSG guideline group will be assessed for coherence with the current review and adaptations made if necessary. The final data extraction form is likely to comprise: publication first author and date; setting; inclusion criteria; population characteristics (age, sex, medications that increase GI bleeding, blood disorders and setting); details relating to the index test and reference standard; and results of key outcomes (diagnostic accuracy metrics, model performance statistics and any additional outcomes as described in Section 2.4). Covariates that may be required for meta-regression will be included in the form, for example, sex and age.

The ACPGBI/BSG team have consented to supply their data extraction files. Their data extractions will be checked against the original publication and checked for completeness against the inclusion criteria for DAP50. Adaptations to the data extraction form and additional data will be extracted where necessary. De novo data will be extracted by one reviewer and checked by a second. Any disagreements will be resolved through discussion and consultation with a third reviewer (e.g., a clinical advisor) where necessary. If time allows, attempts will be made to contact authors for any missing data that are essential to the review. Data from multiple publications of the same study will be extracted as a single study. Where data are reported in other high quality systematic reviews, relevant data may be extracted from the review by one reviewer and checked against the original publication by a second. All studies included by the other reviews will be revisited to ensure any relevant data have not been missed, especially with respect to thresholds and subgroups.

2.9 Quality assessment strategy

The ACPGBI/BSG team have consented to supply their quality assessment files. Their quality assessment decisions will be checked against the original publication. Any disagreements will be resolved through discussion and consultation with a third reviewer (e.g., clinical advisor) where necessary.

For the review of clinical efficacy, the Cochrane Risk of Bias 2[17] tool will be used for RCTs, and the Risk of Bias in Non-Randomised Studies (ROBINS)-I[18] tool will be used for non-randomised clinical trials.

For the review of diagnostic test accuracy, QUADAS-2[19] will be used. For the review of comparative diagnostic test accuracy, QUADAS-C will be used.

2.10 Methods of analysis/synthesis

Where sufficient data exist, pooled estimates of diagnostic parameters will be estimated using a hierarchical meta-analysis model to account for the correlation between sensitivity and specificity.[20-22] If data allows, a single unified analysis that accommodates estimates of sensitivity and specificity at more than one explicit diagnostic threshold per study will be used.[23-25] Tests will be considered individually using meta-regression with test type as a covariate.[26]

If a single unified approach is not feasible, or additional sensitivity analyses are required, separate meta-analyses will be conducted for each FIT assay type, and/or at each explicit threshold reported. Adjustment for imperfect reference standard will also be considered.

Random effects meta-analysis will be used to account for the heterogeneity between studies that is generally expected in diagnostic accuracy studies. Reasons for the heterogeneity in sensitivity and specificity between studies, according to subgroups of interest identified in Section 2.1, may be explored using meta-regression and/or subgroup analyses.

Analyses will be conducted in R[27] using a suitable Markov Chain Monte Carlo (MCMC) sampler such as JAGS[28] or WinBUGS.[29] Results will be displayed as forest plots and summary receiver operating characteristic (SROC) curves with 95% credible intervals (CrI) and 95% prediction intervals (PrI) for sensitivity and specificity.

Statistical synthesis of clinical outcomes will also be conducted if appropriate.

The synthesis plan may be influenced by the requirements of the model with respect to the comparator arm, e.g., if test accuracy data for high-risk NG12 and low-risk DG30 patients is required separately to inform the "current care" pathway (see Section 3.3.2.2).

The impact of the quality of studies on the evidence base may be evaluated through sensitivity analyses in meta-analysis, or through narrative synthesis of the results.

Should data availability and clinical and methodological heterogeneity preclude a meta-analysis, a formal narrative synthesis will be conducted in line with the Synthesis Without Meta-analysis (SWiM) guidelines,[30] and with respect to the subgroups noted in Section 2.1.

2.11 Methods for estimating quality of life – if possible and relevant for the systematic review in question

Quality of life estimates reported within the clinical literature included in this review will be collated as part of the systematic review, whilst data in the cost-effectiveness literature will be identified as part of Section 3.

3. Report methods for synthesising evidence of cost-effectiveness

3.1 Identifying and systematically reviewing published cost-effectiveness studies

A systematic review will be conducted to identify existing cost-effectiveness studies which may be relevant to the decision problem. A literature search will be performed to identify published economic evaluations of the use of FIT in people presenting to primary care with symptoms of CRC. Economic evaluations which are identified by the search which meet the selection criteria for the review will be included. Selection criteria will be developed as part of the review. Data extraction will focus on: (1) the indicated population, main results in terms of costs, consequences and the incremental cost-effectiveness of the alternatives compared, and (2) the modelling methods used, the sources of input parameters, key modelling assumptions and the robustness of the study results. Methodological quality will be assessed using published checklists for economic evaluations and modelling studies.[31, 32]

3.1.1 Search strategy

The following databases and web resources will be searched to identify relevant economic evaluation, utility, and cost studies:

- MEDLINE (Ovid)
- MEDLINE Epub ahead of print, In-Process Citations and Daily Update (Ovid)
- EMBASE (Ovid)
- EconLit (Ovid)
- NHS Economic Evaluation Database (NHS EED) (Internet)
- Tufts' CEA Registry (https://cevr.tuftsmedicalcenter.org/databases/cea-registry)
- Research Papers in Economics (RePEc) (http://repec.org/)

Methodological study type search filters will be applied including the NHS EED filter to identify economic evaluations and other economic filters designed by the McMaster University HEDGES team (https://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx).

3.2 Identifying and reviewing additional published studies to inform the cost-effectiveness analysis

Targeted systematic literature searches will be undertaken to identify studies that can be used to support the development of the health economic model. The studies identified will inform modelling assumptions and estimates of model input parameters. These will aim to inform specific parts of the model such as CRC disease transition probabilities, survival by disease state, costs, morbidity including adverse events associated with COL and utility information that may be used to inform the model parameters. Note that these searches will not constitute a systematic review and will follow the principles of NICE Decision Support Unit (DSU) technical support document 13 (TSD13).[33]

3.3 Health economic model: Evaluation of costs, quality of life and cost effectiveness

A health economic model will be developed to assess the most clinically and cost-effective way to use FIT to detect faecal occult blood as a triage step for the investigation of people presenting to primary care with symptoms or signs indicating risk of CRC (as defined by DG30 and NG12 - the population includes both low and high-risk patients)[2, 3] to reduce the level of referrals to secondary care. The model will explicitly consider the diagnostic accuracy at FIT thresholds and the COL capacity constraints when exploring the relationship between waiting times, time to diagnosis and patient health.

Current COL capacity within the NHS is constrained and there are currently long waiting lists for COL. As described in Section 1.1, in August 2022, 28% of people seen by a specialist for suspected CRC were not seen within 2 weeks of urgent referral, and 53% did not have a diagnosis within 28 days (NHS cancer waiting times, August 2022). Clinical experts also advised that waiting lists for non-urgent referrals to COL are currently much longer than the target 18 weeks. The diagnostic pathways include referrals to lower GI consultations via (1) urgent suspected cancer 2WW referrals and (2) non-priority routine referrals (with an 18WW target). NB, in some areas, non-priority referrals may be further subdivided into a 4WW and an 18WW referral based

on clinical judgement; we will explore the prevalence of this pathway with our advisors. A proportion of each type of referral will go on to receive COL, whilst some will receive other investigations. The investigations received by patients following referrals (such as COL, CTC, CT) may depend on age, COL capacity, regional pathways variations and other patient characteristics; and as both types of referrals will result in people receiving COL, they both contribute to COL capacity requirements.

The decision problem states, "*taking into consideration potential colonoscopy capacity constraints for urgent and non-urgent referrals*". The modelling will define 'referrals' to be all lower GI referrals of symptomatic patients to include both urgent and non-urgent referrals. The modelling will consider a range of referral levels: (1) C, the current referral level, (2) T, a reduction in referrals compared to current referrals which enables target wait times to be met (to be estimated via expert opinion) and (3) a range of referral levels between (1) and (2). Target waiting times being met will be defined according to Operational Standards described in the NHS England Waiting Times for Suspected and Diagnosed Cancer Patients 2020-21 Annual Report of 93% for Two week wait for all cancers and 85% for 62-day wait for first treatment following an urgent GP referral for all cancers. The EAG suggests that there will not be a specific base case (in terms of referral levels).

Different testing strategies which use quantitative FIT in primary care will be assessed (see Section 1.3.2). They will be delivered to all persons with symptoms or signs indicating CRC risk, except for persons with 'bypass symptoms' (see Section 1.9.3). Note that although persons with bypass symptoms are excluded from the decision problem, the size of this population will impact on COL capacity so may be considered to inform the capacity estimates.

The first intervention being assessed, Intervention 1, will be a FIT strategy which uses a single FIT threshold to determine management pathways. A range of different single FIT thresholds will be considered (e.g., 2, 10, 20, 50 and 100 µg Hb/g of faeces) determined by the outputs of the evidence synthesis (see Section 2.10). The second intervention being assessed, Intervention 2, will be a FIT strategy using two thresholds

(t1 and t2, e.g., 10 and 100µg Hb/g) to determine management pathways. A range of different pairs of FIT thresholds will be considered, determined by available data.

The analyses will consider a range of FIT strategies including both intervention 1 and intervention 2 at different FIT thresholds, and with different options for safety netting. Note that the current ACPGBI/BSG recommendation (Intervention 1, threshold of 10µg Hb/g) will be included within this list. Any FIT strategies which would worsen the burden on COL capacity (and therefore may increase average waiting times) will be excluded in line with the scope.

The model will be used to generate predictions for each of these FIT strategies. Predictions will include NHB, referrals levels, impact on waiting times, referrals with no significant bowel pathology and a range of other key outcomes as described in section 1.8. NHB will be calculated assuming a willingness to pay (WTP) of £20,000 per QALY in the base case and a WTP of £30,000 in sensitivity analyses. The FIT strategy with the highest NHB will be identified for a range of referral levels as described above. The results will describe the FIT strategy with the highest NHB for each of the referral levels considered. Other key FIT strategies of interest (such as the ACPGBI/BSG guidelines) would also be included within the results table if they do not increase the level of referrals.

The economic analysis will adopt the perspective of the NHS and PSS. Health outcomes and costs will be evaluated over a lifetime horizon. Modelling assumptions and model parameter values will be taken from published literature (see Section 2, Section 3.2, and Section 3.3.5), study data, routine cost sources, and clinical expert opinion, as required. Costs will be valued at current prices. In line with the NICE Reference Case, health outcomes and costs will be discounted at a rate of 3.5% per year.

3.3.1 Model structure

It is anticipated that the structure of the model developed for this assessment will follow a similar structure to those used in NG12 and DG30, based on a hybrid decision tree / state-transition approach which captures diagnostic pathways, CRC progression, and

CRC and other-cause mortality. The decision tree component of the model has a short time horizon and models the results of investigations for CRC for a cohort of patients presenting to primary care with symptoms which indicate a risk of CRC.

Schematic representations of the anticipated model decision trees are shown in Figure 2 (pathway with one FIT threshold (Intervention 1)) and Figure 3 (pathway with two FIT thresholds (Intervention 2)). The diagnostic and treatment pathway for the intervention strategies are described in Section 3.3.2.1, whilst the diagnostic pathway for current care as a potential comparator is described in Section 3.3.2.2 (the schematic figure for current care is shown in Appendix 2). The decision tree will be followed by a state-transition model with a lifetime time horizon to estimate costs, life years and QALYs for people according to their underlying disease state at diagnosis (Figure 4).

3.3.2 Diagnostic pathways modelled

There is known to be heterogeneity within care pathways across the country. Patient care pathways on which the modelling will focus should be identified via consultation with clinical advisors. We will obtain expert opinion to understand (1) what should happen (according to guidelines and clinical opinion), (2) what the current heterogeneity is in care pathways and (3) what happens in the majority of places. This information will be used to determine what pathways to model for the base case and if any additional pathways should be modelled in scenario analyses if data and time allows.

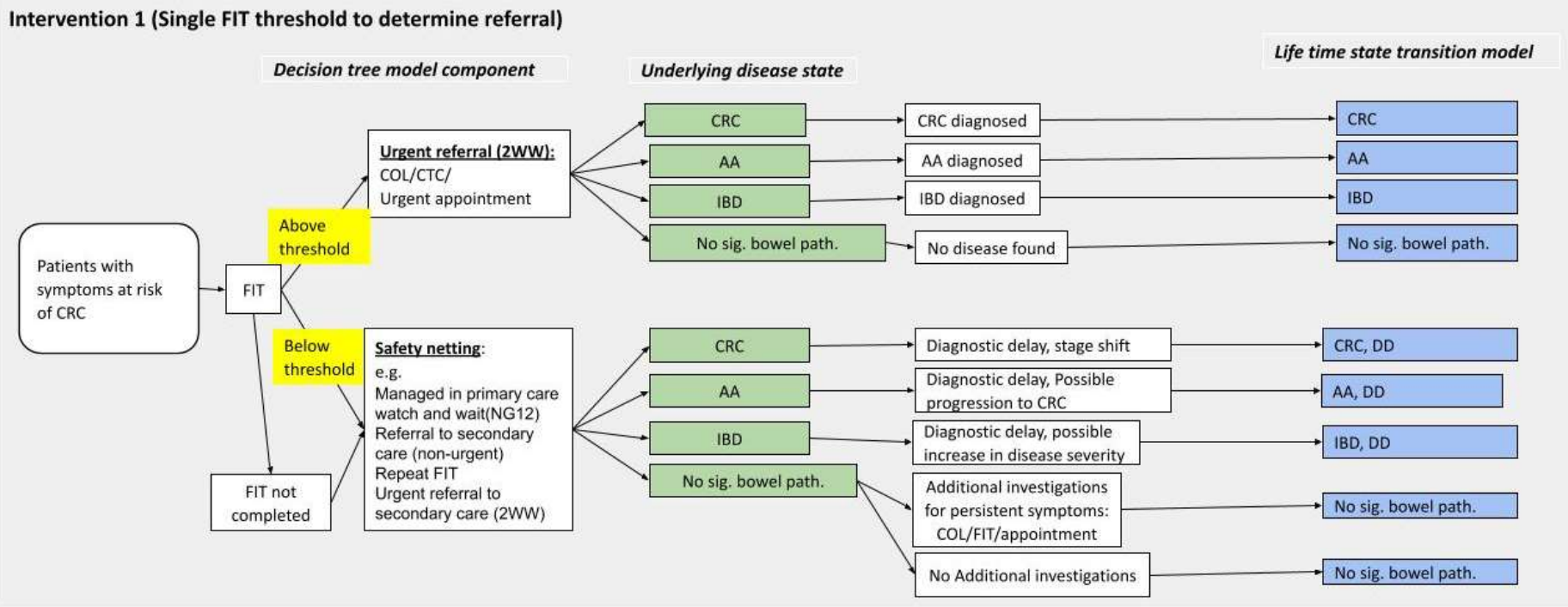
The parts of the pathway which we anticipate seeking expert opinion on will include the following:

(1) Safety netting (see Section 1.9.5): which pathways are included and what proportion of patients follow different safety netting pathways. Potential pathways would be: direct non-urgent referral to secondary care, retesting with FIT, management in primary care (watchful waiting);

(2) For the intermediate risk group (FIT results between t1 and t2), which care pathways would patients follow (see Section 1.9.1.1);

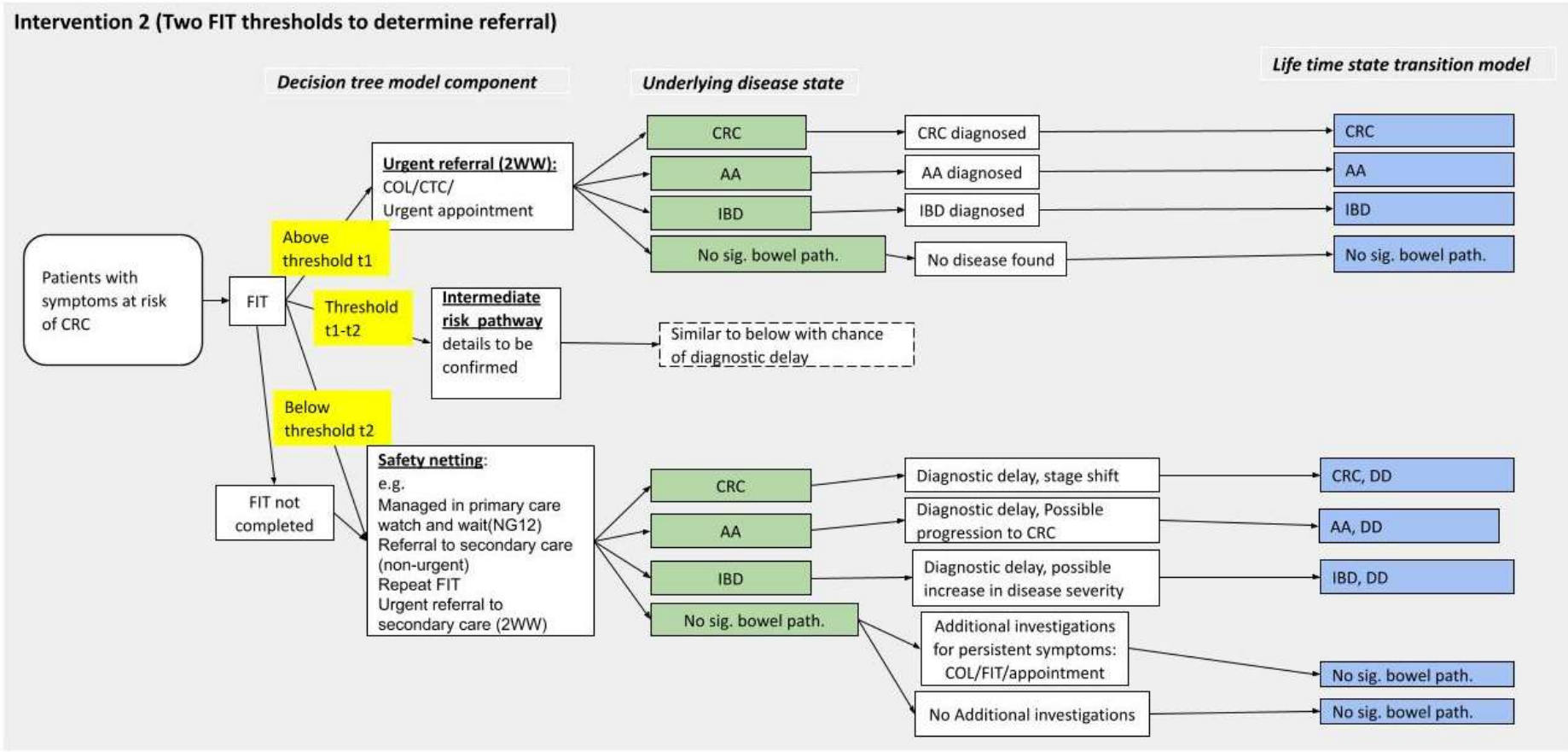
(3) Urgent 2WW suspected CRC referral (see Section 1.2.1 and 1.9.7): which

Figure 2: Anticipated model structure, diagnostic pathway for Intervention 1 (FIT strategy with one threshold)



AA, advanced adenoma; CRC, colorectal cancer; FIT, Faecal Immunochemical Test; IBD, inflammatory bowel disease;2WW, two week wait, DD,delayed diagnosis.

Figure 3: Anticipated model structure , diagnostic pathway for Intervention 2 (FIT strategy with two thresholds)



AA, advanced adenoma; CRC, colorectal cancer; FIT, Faecal Immunochemical Test; IBD, inflammatory bowel disease;2WW, two week wait.

investigations are used and how choice of procedure is determined (e.g., by age: <75 COL, 75-80 CTC, >80 CT; by available COL capacity or other);
(4) Urgent 2WW suspected CRC referral (see Section 1.9.7): whether it includes a telephone triage or an urgent appointment with a consultant or a nurse specialist to determine if COL/CTC is required
(5) Completion of FIT in primary care: which pathways patients who do not complete their FIT would follow.

If current care is included as a comparator, this will reflect the existing NICE guidance available (DG30 and NG12, see description in Sections 1.5 and 3.3.2.2), in accordance with the scope issued by NICE.[34] We note that this may differ from current actual practice in some centres as locations are increasingly moving towards adopting the ACPGBI/BSG guidelines.

The following sections describe the EAG's current understanding of the treatment pathway with and without the proposed interventions. Clinical advice will be sought to ensure these reflect clinical practice.

3.3.2.1 Interventions: FIT in primary care for all patients to guide management

All the model assumptions described in this section and others considered for the scenario analyses will be validated by clinical experts during the development of the model.

Intervention 1: FIT with 1 threshold to determine management

In Intervention 1 (using one FIT threshold) all patients are invited to complete a FIT test by their GP. Patients with a positive FIT result (above the threshold) will be referred by their GP to the suspected cancer pathway in secondary care (urgent 2WW suspected CRC referral). Patients with a negative FIT (below the threshold) are assumed to receive 'safety-netting'.

Intervention 2: Two FIT thresholds guide management

For Intervention 2 (using two thresholds, t1 and t2, see Figure 3) all patients are invited to complete a FIT test by their GP. Patients with a FIT result above the higher threshold (t1) will be referred to the urgent 2WW suspected CRC pathway in secondary care, and patients with FIT results below the lower threshold (t2) will be assumed to receive safety-netting. The follow up for patients with FIT levels between the two thresholds, called the intermediate group here, will be determined via expert opinion.

3.3.2.2 Comparator strategies

For each referral level (C, T, and the range of values in between C and T that are selected for modelling), the comparator will be the FIT thresholds that reduce the number of referrals to COL by at least the amount defined by the referral level. The FIT threshold which provides the most NHB at each given referral level will be indicated. A comparison table providing the maximum NHB at each of the referral levels will be provided to allow exploration of the relationship between reducing the number of referral and the NHB gained. Further details are provided in Section 3.3.3.

The diagnostic pathway for C, current standard care as defined by DG30 and NG12, is described in Section 1.2. In summary, patients, based on their symptomatology and age, can initially receive either: (1) clinical assessment, which leads directly to an urgent 2WW suspected CRC referral pathway (for patients classified as high-risk of CRC by NG12 eligibility criteria); or (ii) a FIT test (for low-risk patients of CRC based on DG30 criteria), which based on a threshold of 10 µg Hb/g can lead to a direct urgent 2WW suspected CRC referral to secondary care or to safety netting. A decision on whether current standard care will be explicitly modelled will be taken during the project. Current care is more complex to model and if it becomes clear that there are larger NHB associated with a reduction in referral levels based on the FIT threshold analyses and clinicians believe that the current method for segregation of patients is inappropriate then the EAG will explore additional sensitivity analyses rather than providing the NHB of current care. The EAG will, however, model FIT strategies that result in current referral levels, as noted above for referral level C.

3.2.2.3 FIT non-completers

For patients who do not complete their FIT, subsequent pathways will be determined via expert opinion (see Section 1.9.9).

3.2.2.4 2WW-referral

The urgent 2WW suspected CRC referral will involve further diagnostic investigation in secondary care (using COL, CTC, or consultation with a specialist). The choice of diagnostic investigation may be dependent on local practice guidelines or other factors (see Sections 1.2.1 and 1.9.7). For example, which investigations are used and how choice of procedure is determined (e.g., by age: <75 years receiving COL, 75-80 years receiving CTC, >80 years receiving CT; by available COL capacity; or other reasons). The proportion of patients receiving each of the investigations will be determined by the data available or via expert opinion, which will also include which groups of patients will go on to receive COL following CTC.

COL can also detect AAs and lower GI pathologies other than CRC (i.e., IBD). The model will also explicitly model these two pathologies (see Section 1.9.6). The model will assume that referral to secondary care (either urgent 2WW suspected CRC referral or 18WW non-urgent referral or other pathways) will result in diagnosis of CRC/IBD/AA if the patient has underlying CRC/IBD/AA i.e., that the diagnostic test (or sequence of tests) used following referral have perfect sensitivity. Therefore, it will be assumed that these patients will go on to receive treatment and will move to the lifetime state transition model.

Patients with an underlying health state of ‘no significant bowel pathology’ will be assumed to have no disease detected at lower-GI referral (i.e., that the diagnostic test (or sequence of tests) used following referral have perfect specificity). These patients will be assumed to move to the lifetime state transition model ‘no significant bowel pathology’.

If time and data allow, the impact of diagnostic tests being an imperfect reference standard may be investigated in the cost effectiveness model.

3.2.2.5 Intermediate group follow-up

The appropriate pathways for patients with FIT levels between the two thresholds (intermediate group, see Section 1.9.1.1) will be based on clinical opinion. In the base-case analysis they are anticipated to receive a non-urgent referral and will be assumed to receive the same investigations in secondary care but less quickly than patients in the urgent 2WW suspected CRC pathway. Scenario analyses exploring variations in the pathways for the intermediate group will be undertaken if data and time allow.

3.2.2.6 Safety Netting Pathways

There is variation in safety-netting available in England (see Section 1.9.5 and 3.3.2).[7, 10, 35] We have considered that safety netting would encompass both the 18WW pathway and watchful waiting where patients may be discharged and may represent to their GP with persistent symptoms.

Patients with underlying CRC, IBD or AA who follow the 18WW pathway will be assumed to receive an accurate diagnosis following the investigation, but which includes a delay due to the wait times on the non-urgent pathway. Depending on the length of this delay, there may be an impact on patient outcomes such as cancer stage progression and QALYs gained.

For persons with underlying CRC/IBD receiving watchful waiting, the model will assume that symptoms will persist, and that these patients will receive a delayed diagnosis. Patients who are not sent to COL but have underlying IBD will still eventually require a COL for their condition to be diagnosed. Hence, a reduction in the number of urgent 2WW suspected CRC referrals will not result in the same reduction in the number of COLs in this group; this impact will be assessed if data and time allow. The likely duration of the delay in diagnosis will be informed by clinical opinion (see Section 3.3.3). For patients without underlying CRC/IBD, the model will assume in the base-case analysis that these patients will be discharged with no significant bowel pathology found.

The EAG will aim to explore scenario analyses which include a more intensive model

of safety netting (see Section 1.9.5), with a proportion of patients in this group receiving additional investigations (e.g., an additional FIT in primary care for patients with an initial negative FIT test used in order to reduce the number of false negatives).

3.3.3 Time to diagnosis/impact of diagnostic delay model component

The model will attempt to estimate the potential benefit of reducing the number of referrals to COL services through the use of different FIT thresholds. The impact of fewer referrals on the time between presentation and diagnosis will be estimated, as will the subsequent impact on patient health (measured in QALYs) although these relationships are likely to be associated with considerable uncertainty. Reductions in the number of referrals should reduce the time to COL for those people who are referred meaning that appropriate treatment can be provided more quickly, increasing QALYs for the cohort. However, patients who are not referred and need appropriate treatment will have a longer delay reducing QALYs for the cohort. Alternative FIT thresholds may allow an increase in total QALYs if the benefits of people receiving prompt treatment outweigh the losses through delayed treatment. More details on estimating the changes in QALYs are provided in Section 3.3.4.

3.3.4 State transition model (life-time horizon) following diagnostic model component

A state transition model will be used to estimate long-term costs and health outcomes (life years and QALYs) for those with a CRC diagnosis, those with other significant lower GI pathology (IBD and AAs, if data allow) and for those who do not have any significant pathologies. Note that the diseases included within the modelling may change depending on clinical opinion on the definition of 'no significant bowel pathologies' and available data.

Patients receiving a diagnosis of CRC will receive a CT scan and/or MRI to establish the stage of the cancer. Patients diagnosed following an urgent 2WW suspected CRC referral will be assumed to experience no delays in diagnosis, and therefore the CRC stage distribution relating to symptomatic diagnosis will be applied. The CRC stage distribution will be estimated based on data for stage distribution for symptomatic CRC

incidence (as opposed to stage distribution for screen detected incidence), which will be potentially based on data from the UK's National Cancer Intelligence Network (NCIN).[36]

Patients with false negative FIT results will be assumed to experience a delay to diagnosis. The length of delay in diagnosis will be informed by available data and/or clinical opinion. The delay in diagnosis may result in these patients being diagnosed at a later stage of the CRC or die (from CRC or another cause). The estimated stage shift (disease progression) associated with the length of diagnostic delay and associated QALY decrement may be estimated using stage transition probabilities for a CRC disease progression model, such as from the MiMiC-Bowel disease progression model.[37, 38]

Once a patient receives a diagnosis of CRC, they are assumed to receive lifetime treatment costs for CRC based on stage at diagnosis. Patients' survival will be modelled conditional on their age and stage at diagnosis.

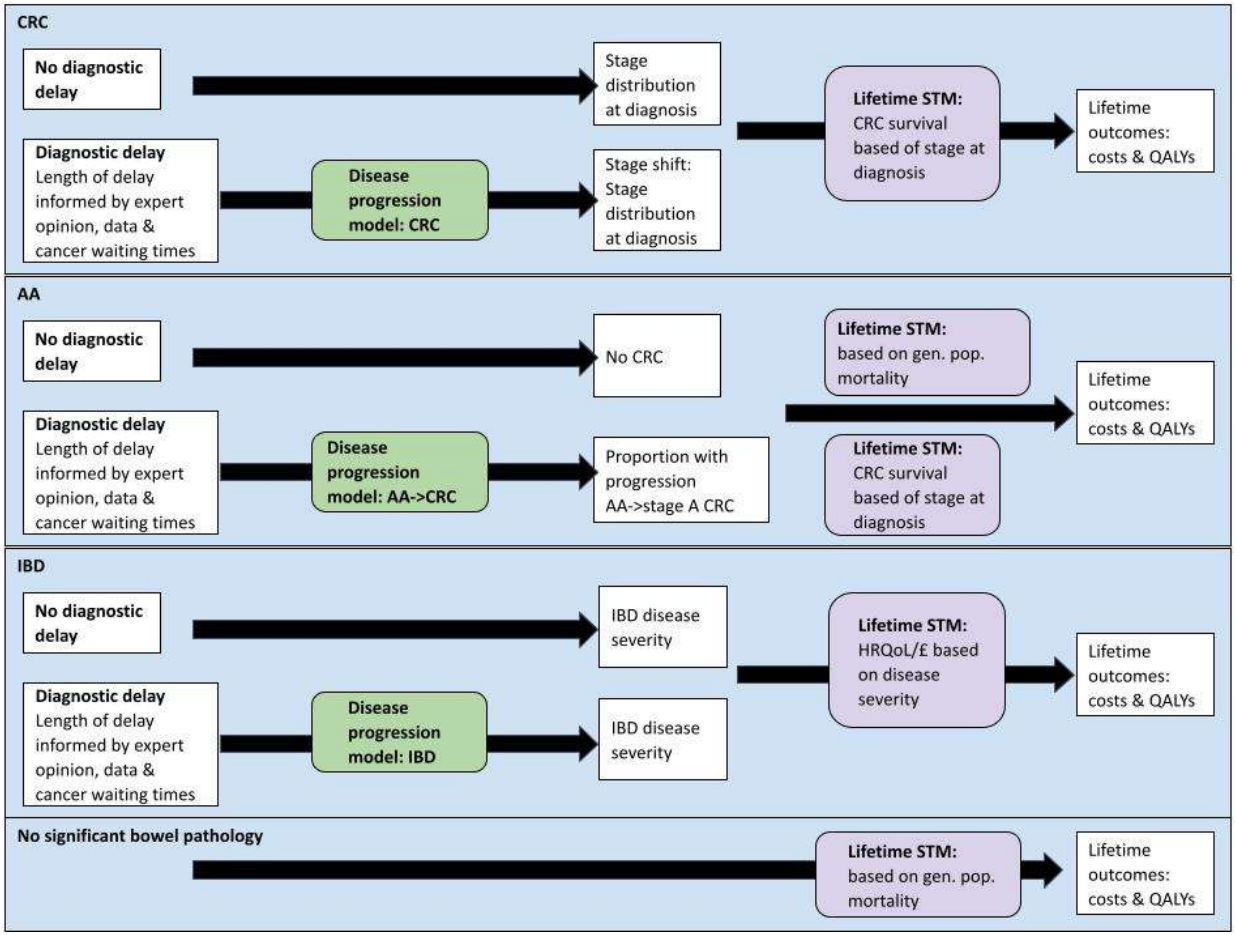
Patients receiving a diagnosis of IBD or AAs will be assumed to receive the follow-up and treatment for these pathologies, and associated costs. Additional assumptions regarding these groups of patients, including their mortality risks, will be based on expert clinical opinion and data from published literature where available.

Survival for individuals without underlying CRC will be modelled using ONS life tables.

3.3.5 Informing model parameters and assumptions

Data highlighted by the systematic review described in Section 2, and by the literature reviews of the existing economic studies and studies to identify health related quality of life (HRQoL), cost and resource use, and other relevant parameters will be considered to inform the modelling (see Section 3.1 and 3.2). Data that most closely matches the decision problem and healthcare context in terms of patient recruitment, ongoing care, location and so on (e.g. UK-based studies such as the NICE FIT study

Figure 4: Anticipated model structure , lifetime state-transition model



AA, advanced adenomas; CRC, colorectal cancer; IBD, inflammatory bowel disease; QALY, quality-adjusted life year; STM, state transition model

and COLOFIT studies [39, 40]) may be prioritised over other sources of data.

It is anticipated that some parameters which will be used to populate the model will be: initial patient characteristics, disease prevalence for CRC and IBD, mortality risks (for general population and patients diagnosed with CRC or IBD), diagnostic HRQoL, and costs related to healthcare management (see Section 3.2.5).

The model is anticipated to include resource costs associated with: (i) use of FIT in primary care; (ii) imaging tests in secondary care; (iii) clinical appointments and resource use related to primary care and referrals to secondary care; (iv) resource use related to staging CRC; (iv) lifetime CRC treatment costs by health state; (v) management of other-than-CRC low GI pathologies detected by COL/CTC; (vi) management of adverse events associated with COL, and (vii) management of further diagnostic work up, including those associated with false test results and subsequent inappropriate treatment.

Resource costs associated with FIT kits will include: cost of equipment, reagents and consumables, and staff and associated training costs where data allow. Resource use related to primary care and referrals to secondary care will include clinical appointments with GPs, nurses and consultants, the costs of investigations (such as COL, CTC, and CT), and the use of further tests (e.g., blood tests and imaging tests used for CRC staging). Costs incurred from adverse events associated with various complications from COL, such as bleeding and perforation of intestine, will be included in the model and will be assumed to be resolved without further long-term costs. Lifetime CRC (by CRC stage), AA and IBD treatment costs will be included in the model. Other types of costs, such as the costs of repeat and dual FIT tests, those related to the use of other resources in primary and secondary care (e.g., e-referral systems) and costs associated with false test results, which result in further diagnostic work up and subsequent inappropriate treatment may be included in the model if data allow.

Resource costs will be valued using unit costs obtained from routine costing sources (e.g. NHS Reference Costs,[41] the PSSRU,[42] the BNF[43]), existing studies retrieved by

the literature review, and through personal communication with relevant clinical experts and technology manufacturers.

HRQoL values will be applied to each of the model health states and will be based on literature and/or other available sources. Utility decrements related to adverse events associated with COLs and patient-reported outcomes for anxiety associated with waiting times for referral to secondary care and test results due to diagnosis delays, and further diagnostic workup will be considered for inclusion if evidence permits. QALYs will be calculated from the economic modelling, by multiplying the life years that patients spend in each health state of the model by the associated utility, representing the valuation of the health state of the patient and considering the impact of transient events. Utility values will be age-adjusted using Hernández Alava *et al* (2022).[44]

3.3.6 Model Analyses

Section 3.3 describes the main analyses that will be undertaken in this assessment to estimate the reduction in referral levels and the FIT threshold which is most cost-effective. Reporting of the economic analysis will follow the CHEERS checklist.[45].

Probabilistic sensitivity analysis (PSA) will be run and estimates of cost-effectiveness will be based on the results of the PSA. Value of Information Analysis will be undertaken if time allows. Deterministic sensitivity analyses (DSA) will be performed to identify key drivers of the cost-effectiveness of FIT.

In addition, the EAG will seek to undertake the following scenario analyses to explore the impact of key modelling assumptions:

- Using a WTP £30,000 per QALY
- Model analyses considering different FIT assays and analysers (see Section 1.9.2)
- Dual testing: the use of dual FIT testing (where two samples from different bowel movements are collected rather than a single sample for the same initial FIT test) (see Section 1.9.4).

- Scenarios exploring different assumptions around the proportion of patients receiving different pathways with safety netting. To include safety netting pathways reported by the North London or Oxford groups (see Section 1.9.5), and different assumptions about the intermediate group pathway (see Section 1.9.1.1)
- Scenarios in which AAs are not included (see Section 1.9.6)
- Analyses relating to subgroups of the population:
    - that may require a different threshold of FIT (see Section 1.9.1), such as age, sex, ethnicity, people taking medications which increase the risk of gastrointestinal bleeding, people with blood disorders that could affect the performance of the test, and people with anaemia (including iron deficiency anaemia).
    - who may be associated with different levels of compliance with FIT, e.g., by socioeconomic status or ethnicity.
    - NB a decision will be made about whether it is useful to model these subgroups based on a) the evidence found by the clinical review relating to the need for an alternative threshold in these groups, or differential compliance with FIT, and b) an assessment of whether this would cause equity issues.
- Adjustment for imperfect reference standard and/or use of other reference standards if time and data allow

4. Handling information from the companies

All data submitted by the manufacturers/sponsors will be considered if received by the EAG no later than 1st February 2023. Data arriving after this date will not be considered, unless it was specifically requested by the EAG, and a later date of submission agreed. If the data meet the inclusion criteria for the review, they will be extracted, and quality assessed in accordance with the procedures outlined in this protocol.

Any 'commercial in confidence' data provided by a manufacturer and specified as such will be highlighted in blue and underlined in the assessment report (followed by an indication of the relevant company name e.g., in brackets). Any 'academic in confidence' data provided by the manufacturer, and specified as such, will be highlighted in yellow and underlined in the assessment report. Any confidential data used in the cost-effectiveness model will also be highlighted. Any fully incremental analyses which include CIC data from multiple manufacturers will be highlighted in green and underlined. If confidential information is included in economic models, then a version using dummy data or publicly available data in place of confidential data will be provided.

5. Competing interests of authors

There are no competing interests for Sue Harnan, Aline Nevaga Biz, Sophie Whyte, Katy Cooper, Jean Hamilton, Kate Ren, Mark Clowes, Andrea Shippam, Matt Stevenson, and Louise Merriman.

Willie Hamilton is the Medical Director of a mutual friendly society that offers health insurance, but his pay is not linked to performance of the company. He is also Co-PI on an NIHR HTA award, which seeks to optimise the use of FIT in colorectal cancer diagnosis. Stephanie Edgar has no financial conflicts but is employed by Cancer Research UK who may have an interest in the outcome of the appraisal. Alex Ball and Matt Kurien have no financial conflicts, but their research group has recently published a study examining the diagnostic accuracy of the faecal immunochemical test in low-risk symptomatic populations. Kevin Monahan, Richard Booth, Rachel Carten and Muti Abulafi are co-authors of the ACPGBI/BSG guideline (KM, RB, RC, MA) and associated systematic review (RB, RC, MA).

6. Timetable/milestones

Milestone	Date to be completed
Final date for Manufacturer/sponsor data submissions	1st February 2023
Progress Report	10th February 2023
Draft Assessment Report	20th April 2023
Final Report to NICE	19th May 2023

7. Appendices

Appendix 1: Example search strategy (clinical effectiveness)

Ovid MEDLINE(R) ALL <1946 to December 02, 2022>

1 f?ecal immunochemical test.mp. 1256
2 f?ecal occult blood.mp. 4445
3 f?ecal h?emoglobin.mp. 269
4 ((immunochromatographic or immuno-chromatographic or immunochem$ or immuno-chem$ or immunohistochem$ or immuno-histochem$ or immunol$ or immunoassay or immuno* assay or immunoturbidimetric or immunosorbent or elisa) adj4 (f?ecal or f?eces or stool or stools or FIT)).mp. 3593
5 (iFOBT or qFIT).mp. 208
6 or/1-5 7284
7 F?ecal h?emoglobin.ti,ab,ot,hw. 256
8 H?emoccult.ti,ab,ot,hw. 728
9 FOBT.ti,ab,ot,hw. 1429
10 7 or 8 or 9 2335
11 (f?ecal or f?eces or stool or stools).ti,ab,ot,hw. 211763
12 occult blood/ or occult blood.ti,ab,ot,hw. 8920
13 (test$ or measur$ or screen$ or exam$).ti,ab,ot,hw. 10697959
14 11 and 12 and 13 6020
15 6 or 10 or 14 8730
16 exp colorectal neoplasms/ 231063
17 exp cecal neoplasms/ 6040
18 ((colorect$ or rectal$ or rectum$ or colon$ or sigma$ or sigmo$ or rectosigm$ or bowel$ or anal or anus) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot,hw. 323802
19 CRC.ti,ab,ot. 43375
20 ((cecum or cecal or caecum or caecal or il?eoc?ecal or il?eoc?ecum) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 2753
21 (large intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or

carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 1839

22 (lower intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 34

23 16 or 17 or 18 or 19 or 20 or 21 or 22 335647

24 15 and 23 5934

25 limit 24 to yr="2022 -Current" 423

26 (FOB gold$ or FOBgold$ or SENTiFIT).mp. 38

27 (JACK-arc$ or JACKarc$ or HM-JACK$ or HM JACK$ or HMJACK$).mp. 23

28 (OC Sensor$ or OC-Sensor$ or OCSensor$ or [redacted]).mp. 370

29 (OC Pledia$ or OC-Pledia$ or OCPledia or OC-iO).mp. 0

30 (NS-Prime or NSPrime or NS-Plus).mp. 37

31 (POC FIT QRG or POCFITQRG).mp. 0

32 (immundiagnostik or IDK or turbifit or turbitube).mp. 125

33 quikread.mp. 19

34 or/25-33 991

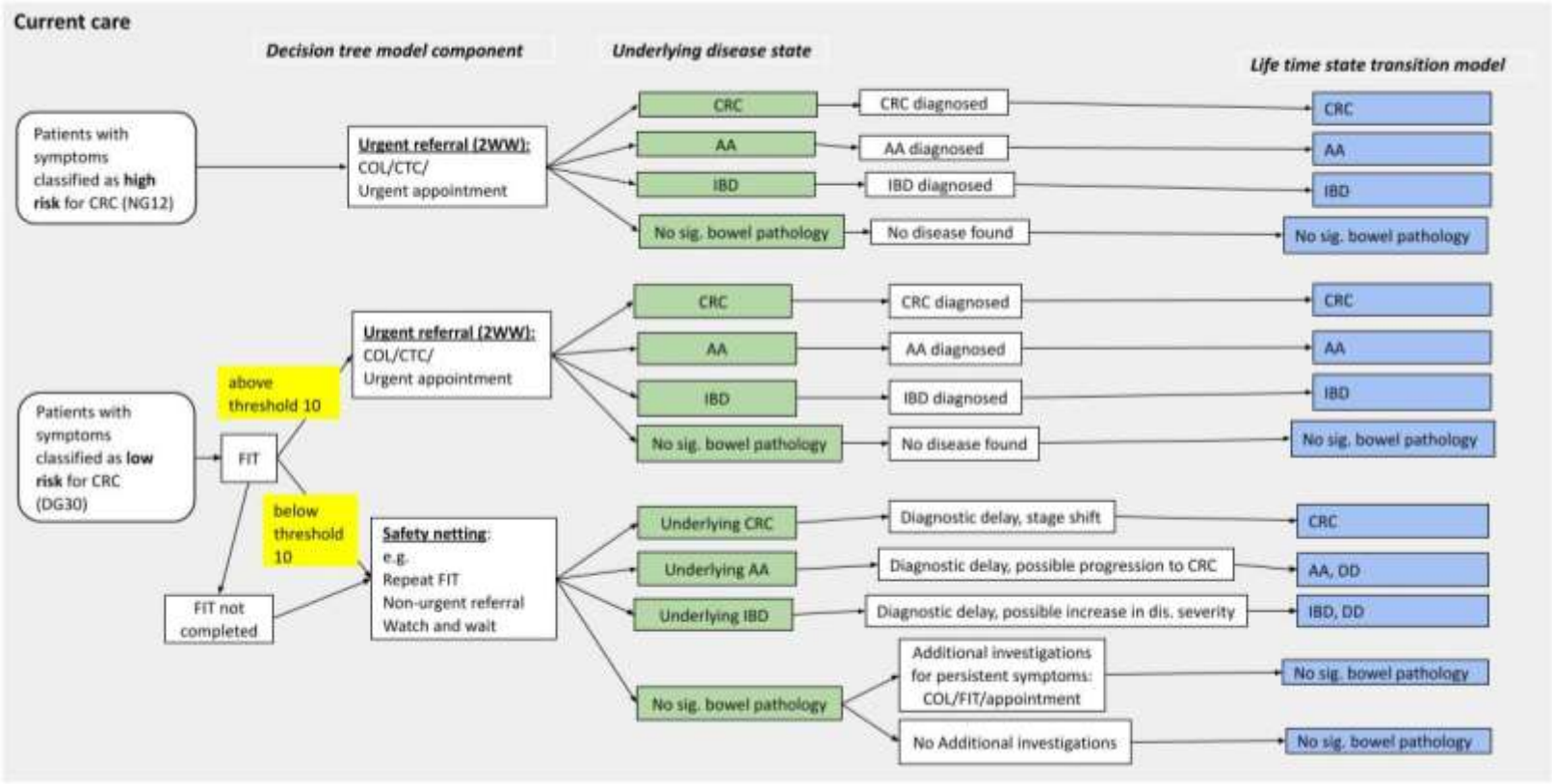
35 limit 34 to yr="2016 -Current" 737

36 exp animals/ not (exp animals/ and humans/) 5070893

37 35 not 36 726

Appendix 2: Current care anticipated model diagnostic pathway structure

Figure 4: Anticipated model structure, diagnostic pathway - current care



AA, advanced adenoma; CRC, colorectal cancer; FIT, Faecal Immunochemical Test; IBD, inflammatory bowel disease; DG30, Diagnostic Guidance 30; NG12, NICE Guideline 12; 2WW, two week wait.

Additional information that is needed by NETSCC, HTA and NICE.

Please send this as a WORD document when you submit your protocol to Htatar@soton.ac.uk.

Details of EAG

Please give details of members of the team

Ms Sue Harnan
Senior Research Fellow in Systematic Reviews
ScHARR, University of Sheffield
Regent Court
30 Regent Street
Sheffield
S1 4DA
Telephone number: +44 (0)114 2220869
Email address: s.harnan@sheffield.ac.uk

Dr Aline Navega Biz
Research Associate in Economic Modelling
ScHARR, University of Sheffield
0114 222 6300
a.navega-biz@sheffield.ac.uk

Dr Jean Hamilton
Research Fellow in Statistics
ScHARR, University of Sheffield
+44 (0)114 2225447
jean.hamilton@sheffield.ac.uk

Dr Sophie Whyte
Senior Research Fellow in Economic Modelling
ScHARR, University of Sheffield
+44 (0)114 222 0859

sophie.whyte@sheffield.ac.uk

Dr Katy Cooper
Senior Research Fellow in Systematic Reviews
ScHARR, University of Sheffield
+44 (0)114 222 0773
k.l.cooper@sheffield.ac.uk

Mr Mark Clowes
Information Specialist
ScHARR, University of Sheffield
+44 (0)114 215 9125
m.clowes@sheffield.ac.uk

Dr Kate (Shijie) Ren
Senior Research Fellow in Statistics
ScHARR, University of Sheffield
+44 (0)114 222 0696
s.ren@sheffield.ac.uk

Professor Matt Stevenson
Professor of Health Economic Modelling
ScHARR, University of Sheffield
0114 222 0691
m.d.stevenson@sheffield.ac.uk

Ms Andrea Shippam
Programme Manager
ScHARR, University of Sheffield
+44 (0)114 222 0693
a.shippam@sheffield.ac.uk

Please indicate to whom you wish all correspondence to be addressed

Please send all correspondence to Sue Harnan (project lead and lead systematic reviewer), Aline Navega Biz (health economics lead), Sophie Whyte (health economics), Matt Stevenson (health economics), Jean Hamilton (lead statistician), Katy Cooper (systematic reviewer), Kate Ren (statistics) and Andrea Shippam (project manager).

REFERENCES

1. Cancer Research UK. Bowel cancer statistics. 2022. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer (Accessed
2. NICE. Quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care. Diagnostics guidance [DG30]; 2017.
3. NICE. Suspected cancer: recognition and referral. 2015.
4. D'Souza N, Abulafi M. The current state of use of the Faecal Immunochemical Test (FIT) in symptomatic patients in the UK: a survey of ACPGBI members. *Colorectal Disease: the Official Journal of the Association of Coloproctology of Great Britain and Ireland* 2022.
5. Abulafi M, D'Souza N. The state of the faecal immunochemical test in symptomatic patients in the UK. *Ann R Coll Surg Engl* 2022;104:240-1.
6. Mozdiak E, Weldeselassie Y, McFarlane M, Tabuso M, Widlak MM, Dunlop A*, et al.* Systematic review with meta-analysis of over 90 000 patients. Does fast-track review diagnose colorectal cancer earlier? *Alimentary pharmacology & therapeutics* 2019;50:348-72.
7. Monahan KJ, Davies MM, Abulafi M, Banerjea A, Nicholson BD, Arasaradnam R*, et al.* Faecal immunochemical testing (FIT) in patients with signs or symptoms of suspected colorectal cancer (CRC): a joint guideline from the Association of Coloproctology of Great Britain and Ireland (ACPGBI) and the British Society of Gastroenterology (BSG). *Gut* 2022;71:1939.
8. Booth R, Carten R, D'Souza N, Westwood M, Kleijnen J, Abulafi M. Role of the faecal immunochemical test in patients with risk-stratified suspected colorectal cancer symptoms: A systematic review and meta-analysis to inform the ACPGBI/BSG guidelines. *The Lancet Regional Health-Europe* 2022;23:100518.
9. D'Souza N, Hicks G, Benton S, Abulafi M. The diagnostic accuracy of the faecal immunochemical test for colorectal cancer in risk-stratified symptomatic patients. *The Annals of The Royal College of Surgeons of England* 2020;102:174-9.
10. . Using Faecal Immunochemical Testing (FIT) in the Lower Gastrointestinal (GI) pathway (primary care). London; 2022.
11. . Using Faecal Immunochemical Testing (FIT) in the Lower Gastrointestinal (GI) pathway (secondary care). London; 2022.
12. Fraser CG, Allison JE, Young GP, Halloran SP, Seaman HE. Improving the reporting of evaluations of faecal immunochemical tests for haemoglobin. *European Journal of Cancer Prevention* 2015;24:24-6.
13. NICE. Quantitative faecal immunochemical tests to guide referral in primary care for people with high risk symptoms. Final scope. 2020.
14. Westwood M, Corro Ramos I, Lang S, Luyendijk M, Zaim R, Stirk L*, et al.* Faecal immunochemical tests to triage patients with lower abdominal symptoms for suspected colorectal cancer referrals in primary care: A systematic review and costeffectiveness analysis. 2016.
15. Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in health care. 2008.
16. The Cochrane Handbook. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). 2019.
17. Sterne JA, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I*, et al.* RoB 2: a revised tool for assessing risk of bias in randomised trials. 2019;366.
18. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M*, et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. 2016;355:i4919.

19. Whiting P, Rutjes A, Westwood M, Mallett S, Deeks J, Reitsma J*, et al.* QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011;155:529-36.
20. Chu H, Cole S. Bivariate meta-analysis of sensitivity and specificity with sparse data: a generalized linear mixed model approach. *J Clin Epidemiol* 2006;59:1331-2.
21. Reitsma J, Glas A, Rutjes A, Scholten R, Bossuyt P, Zwinderman A. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982-90.
22. Rutter C, Gatsonis C. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Statist Med* 2001;20:2865-84.
23. Jones HE, Gatsonsis CA, Trikalinos TA, Welton NJ, Ades AE. Quantifying how diagnostic test accuracy depends on threshold in a meta-analysis. *Statistics in Medicine* 2019;38:4789-803.
24. Menten J, Lesaffre E. A general framework for comparative Bayesian meta-analysis of diagnostic studies. *BMC Medical Research Methodology* 2015;15:70.
25. Steinhauser S, Schumacher M, Rücker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. *BMC medical research methodology* 2016;16:97.
26. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Chapter 10: analysing and presenting results. London, UK: The Cochrane Collaboration; 2010.
27. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing,. In. Vienna, Austria; 2019.
28. Plummer M. rjags: Bayesian Graphical Models using MCMC. R package version 4-10 In; 2019.
29. Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Statistics in medicine* 2009;28:3049-67.
30. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S*, et al.* Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. 2020;368:l6890.
31. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C*, et al.* Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices--Modeling Studies. *Value Health* 2003;6:9-17.
32. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the Economic Evaluation of Health Care Programmes; 2005.
33. Kaltenthaler E, Tappenden P, Paisley S, Squires H. NICE DSU Technical Support Document 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models. In: Sheffield Uo, ed. DECISION SUPPORT UNIT. Sheffield, UK; 2011: 72.
34. NICE. Quantitative faecal immunochemical tests to guide colorectal cancer pathway referral in primary care. Final scope. 2022.
35. NHS North Central London. Faecal Immunochemical Test (FIT). 2022. https://gps.northcentrallondon.icb.nhs.uk/service/faecal-immunochemical-test-fit#overview (Accessed
36. Network NCI. The National Cancer Registration and Analysis Service. In; 2019.
37. Thomas C, Mandrik O, Whyte S. Development of the Microsimulation Model in Cancer of the Bowel (MiMiC-Bowel), an Individual Patient Simulation Model for Investigation of the Cost-effectiveness of Personalised Screening and Surveillance Strategies. 2020.
38. Mandrik O, Thomas C, Strong M, Whyte S. Calibration and Validation of the Microsimulation Model in Cancer of the Bowel (MiMiC-Bowel), an Individual Patient Simulation Model for Investigation of the Cost-effectiveness of Personalised Screening

and Surveillance Strategies. *HEDS Discussion Paper Series* 2021.
39. NHS England. THE NICE FIT STUDY. 2019.
40. (NIHR) NIfHaCR. COLOFIT - Optimal use of Faecal Immunochemical Testing for patients with symptoms of possible colorectal cancer.; 2021. https://fundingawards.nihr.ac.uk/award/NIHR133852 (Accessed
41. NHS Improvement. National Cost Collection: National Schedule of NHS costs - Year 2020-21. 2021.
42. Jones; K, Burns; A. Unit Costs of Health and Social Care 2021. In. Personal Social Services Research Unit, University of Kent, Canterbury; 2021.
43. British National Formulary. MedicinesComplete. In. London: Pharmaceutical Press.; 2022.
44. M Hernández Alava, S Pudney, A Wailoo. Estimating EQ-5D by age and sex for the UK. 2022.
45. Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C*, et al.* Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. *BMJ* 2022;376:e067975.