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)


ABSTRACT
Faecal immunochemical testing (FIT) has a high sensitivity for the detection of colorectal cancer (CRC). In a symptomatic population FIT may identify those patients who require colorectal investigation with the highest priority. FIT offers considerable advantages over the use of symptoms alone, as an objective measure of risk with a vastly superior positive predictive value for CRC, while conversely identifying a truly low risk cohort of patients. The aim of this guideline was to provide a clear strategy for the use of FIT in the diagnostic pathway of people with signs or symptoms of a suspected diagnosis of CRC. The guideline was jointly developed by the Association of Coloproctology of Great Britain and Ireland/British Society of Gastroenterology, specifically by a 21-member multidisciplinary guideline development group (GDG).

A systematic review of 13 525 publications was undertaken to develop 23 evidence and expert opinion-based recommendations for the triage of people with symptoms of a suspected CRC diagnosis in primary care. In order to achieve consensus among a broad group of key stakeholders, we completed an extended Delphi of the GDG, and also 61 other individuals across the UK and Ireland, including by members of the public, charities and primary and secondary care. Seventeen research recommendations were also prioritised to inform clinical management.

OBJECTIVE
To provide a clear strategy for the use of faecal immunochemical testing (FIT) in the diagnostic pathway of people with signs or symptoms of a suspected diagnosis of colorectal cancer (CRC).

BACKGROUND
Evaluation in primary care of symptomatic patients with a potential diagnosis of CRC is challenging. Symptoms alone are unreliable predictors of those who may have a diagnosis of CRC and may therefore result in a high proportion of eligible patients not having access to diagnostic examination. Use of FIT offers considerable advantages over the use of symptoms, with a vastly superior positive predictive value (PPV) for CRC, while conversely identifying a truly low risk cohort of patients. FIT provides an opportunity to effectively triage patients with bowel symptoms into two groups: those who require ‘Fast Track’ referral on an urgent suspected cancer pathway and lower risk patients who may potentially be managed in primary care. The benefit of this stratification should be to reduce the fear of missed/delayed diagnosis of CRC which is currently driving high referral rates for investigation, enabling more effective use of investigative processes with a focus on evaluating those with a significant risk of an underlying CRC diagnosis.

Through the COVID-19 pandemic, FIT has been increasingly employed across the UK, in an ad hoc way in primary and secondary care, leading to significant variation in practice. The purpose of this guideline is to provide an evidence-based framework for the optimal use of FIT in the diagnostic pathway for people with symptoms or signs of a suspected diagnosis of CRC.

METHODS
This guideline was jointly commissioned by the Association of Coloproctology of Great Britain and Ireland (ACPGBI), and British Society of Gastroenterology (BSG), and a guideline chair selected from each society (MMD and KJM). It was developed in accordance with the BSG National Institute of Health and Care Excellence (NICE)-accredited guideline process.

The guideline development group (GDG) included colorectal surgeons (MMD, MA, AB, MM and RJC), nurse specialist (MP) and gastroenterologists (RA, JEE and KJM) nominated by ACPGBI and BSG (co-leads MMD and KJM), general practitioners (GPs: BDN and LSa), a professional

scope

A scoping meeting was held on 20th May 2021, and in advance of this meeting the GDG was asked to develop key priorities and questions. It was agreed that the scope of this guideline was to develop guidelines for the role of FIT testing in patients with signs and symptoms of a suspected diagnosis of CRC. The target audience are clinicians involved in this pathway from primary care through to secondary care.

FIT has a high sensitivity and PPV compared with the use of symptoms alone to determine the need for referral from primary care to secondary care for further diagnostic investigation for people with symptoms or signs of a suspected diagnosis of CRC. The sensitivity of FIT for what is known as other ‘serious bowel disease’ (including advanced adenomas, inflammatory bowel disease) is considerably lower than for CRC, and as such this guideline focuses on the role of FIT in the diagnostic pathways for CRC. Similarly, this guideline is not designed to provide advice for the investigation management of gastrointestinal (GI) symptoms outside the context of FIT in the diagnostic pathway for suspected CRC.

The GDG agreed that these guidelines should stimulate greater efforts to ensure access to FIT for all GPs and eligible patients, and should therefore include advice designed to facilitate implementation.

The GDG developed key questions for the guideline:
1. What FIT thresholds should be used to trigger referral from primary care?
2. Should FIT be used in primary care or secondary care?
3. What advice can we offer clinicians where patients have not returned a FIT test?
4. What safety netting strategies may be employed to avoid missed CRC diagnosis in patients with a FIT below the chosen threshold?
5. What is the diagnostic accuracy of FIT for CRC with specific symptoms?
6. Does diagnostic accuracy vary by patient-related factors (eg, age-group, sex, ethnicity and deprivation)?
7. Is a repeat/second FIT useful and does it enhance diagnostic accuracy?
8. Does the diagnostic accuracy of FIT vary with the type of analyser used?
9. Should FIT be combined with other factors to optimise risk stratification?
10. Can FIT be used in specific populations, for example, young symptomatic patients to facilitate early diagnosis of early onset CRC?
11. Is there a role for specific interventions according to patient or test related factors? Can FIT (faecal haemoglobin (fHb) levels) be used to prioritise investigations?
12. What is the acceptability of FIT in patients with suspected CRC symptoms and their treating clinicians?
13. How can we avoid discriminating against certain populations in this guideline?
14. What lessons may be learnt from implementation programmes of FIT in symptomatic populations?

The GDG agreed that FIT should not be a sole arbiter of referral. Therefore patients who do not have signs or symptoms of a suspected diagnosis of CRC have not been considered within this guideline, and furthermore should not be referred from primary care outside the context of national screening programmes because of a FIT test alone. Conversely those with a FIT below threshold may be managed in primary care, but should not be denied access from referral to secondary care if referral is appropriate for other reasons. Such patients may be referred on routine or urgent pathways, but not necessarily on the suspected CRC investigation pathway. Patients with signs of an abdominal mass should be referred urgently, however a FIT should be requested simultaneously in primary care in order to inform subsequent investigation. Those with an anal/rectal mass or anal ulceration should be referred urgently from primary care without a FIT.

Picos, search strategy and grade

Three broad PICOs (patients, interventions, controls and outcomes) were developed which considered these questions (online supplemental file 1). The Appraisal of Guidelines for Research and Evaluation (AGREE II) instrument provided a methodological framework.

Search strategies agreed by the GDG, and a systematic literature search was performed by four members of the GDG (MA, ND, RVC and RB), which returned 100 publications. Additional references were obtained by cross-referencing and by recommendation from the GDG. Relevant published national and international guidelines were also scrutinised. After each round of Delphi, and before the guideline was finalised, the search was repeated, and any important studies published since the initial evidence search incorporated.

A modified electronic Delphi process was used to develop and refine statements. Initial draft statements formulated by the writing committee were reviewed by the GDG to allow for modification and to identify additional references. After a preliminary discussion, formal anonymous voting rounds were undertaken using SurveyMonkey. Each statement was scored by each member of the GDG using a 5-point Likert scale. We invited also key national and international opinion leaders from the ACPGBI, BSG, primary care, clinical biochemistry, patient representation (who contributed to the online supplemental Lay Summary) and CRC charities to participate in the modified Delphi process. We included additional patient and public involvement in the Delphi process by inviting participants through the national charities Bowel Cancer UK and Bowel Research UK. Consensus required at least 80% agreement. Where consensus was not reached, feedback from the GDG members was disseminated after each round to allow members to reconsider their original position. Where appropriate, revisions to statements were made and a further voting round was undertaken in second and third rounds, with a summary flowchart agreed by the GDG (figure 1).

The GDG also developed research recommendations (online supplemental file 2) which were prioritised by electronic voting. The Grading of Recommendations, Assessment, Development and Evaluations (GRADE) tool was used to evaluate the strength of evidence and the strength of recommendations made (see Executive summary of recommendations). The GRADE system specifically separates the strength of evidence from the strength
1. We recommend that FIT should be used by primary care clinicians to prioritise patients with clinical features of CRC for referral for urgent investigation
2. We recommend that a FIT threshold of fHb ≥10 μg Hb/g should be used in primary care to select patients with lower GI symptoms for an urgent referral pathway for CRC investigation.
3. We recommend that patients should not be excluded from referral from primary care for symptoms on the basis of FIT testing alone.
4. We suggest that clinicians should follow-up patients with no FIT result to encourage them to return a sample or, where the kit has been lost or inadequately submitted, offer a further test.
5. We suggest that patients who decline to return a FIT test should be counselled that evaluation of their symptoms is incomplete, and be encouraged to complete their test.
6. We suggest that where no FIT result can be obtained, clinicians should use existing national and local guidelines to assess risk of CRC.

Safety netting
8. We recommend that some patients with symptoms of suspected CRC may be managed in primary care if fHb < 10 μg Hb/g, and provided appropriate safety-netting is in place.
9. We suggest that patients with an fHb < 10 μg Hb/g but with persistent and unexplained symptoms for whom the GP has ongoing clinical concern should be referred to secondary care for evaluation.
10. We recommend that safety-netting protocols should incorporate advice and strategies for the diagnosis of CRC and extracolonic cancer, as well as other serious gastro-intestinal conditions.

Diagnostic accuracy of FIT for CRC with suspected cancer signs or symptoms
11. FIT is a triage tool to identify those patients with symptoms of suspected CRC who should undergo further colorectal investigation.
12. We suggest that FIT be used for people with iron deficiency anaemia within primary care to inform urgency of referral.
13. We suggest referral of patients with persistent/recurrent anorectal bleeding for flexible sigmoidoscopy if fHb < 10 μg Hb/g.
14. There is currently insufficient evidence to recommend variations in the fHb threshold for referral from primary care according to patient-related factors.

15. There is currently insufficient evidence to confirm whether diagnostic accuracy is impacted by the type of FIT analyser used.

16. There is currently insufficient evidence to recommend including FIT in a risk score with other clinical features to identify patients with symptoms of suspected CRC.

17. We suggest that FIT may be used to stratify adult patients aged younger than 50 years with bowel symptoms suspicious of a diagnosis of CRC.

Investigation in secondary care

18. Colonoscopy is considered the standard method of investigation, however other methods of colorectal imaging may be appropriate in some patients.

19. We recommend that for patients with symptoms of a suspected diagnosis of CRC, CT colonography (CTC) is equivalent to colonoscopy for detection of CRC (the choice of modality should be determined by the local expertise and availability).

20. There is currently insufficient evidence to support use of a specific quantitative FIT threshold to recommend the selection of CT colonography versus colonoscopy.

Acceptability

21. On the basis of limited evidence, clinicians and patients consider FIT as an acceptable test for symptomatic CRC in most circumstances.

22. We recommend that services should consider ways of promoting a high proportion of patients to return FIT kits.

Discrimination

23. We recommend that clinicians actively prevent discrimination at any stage of the diagnostic pathway as symptomatic FIT testing is rolled out, with a focus on equity of access and application to all patients with lower GI symptoms.

Implementation

24. We recommend that FIT, as a diagnostic triage tool, can be implemented safely at primary care level, and that a programme of education be developed to facilitate implementation of FIT in primary care.

FIT IN PRIMARY CARE

We recommend that FIT should be used by primary care clinicians to prioritise patients with clinical features of CRC for referral for urgent investigation:

GRADE of evidence: low; Strength of recommendation: Strong.

We recommend that a FIT threshold of fHb ≥10 µg Hb/g should be used in primary care to select patients with lower GI symptoms for an urgent referral pathway for CRC investigation:

GRADE of evidence: low; Strength of recommendation: Strong.

We recommend that patients should not be excluded from referral from primary care for symptoms on the basis of FIT testing alone:

GRADE of evidence: very low; Strength of recommendation: Strong.

A FIT ≥10 µg Hb/g faeces is recommended by NICE to select patients for an urgent ‘2-week-wait’ (2WW) referral for CRC investigation.5 This recommendation has not changed since it was introduced in 2017 despite the guidance changing on which patient groups, symptoms, signs or anaemias should trigger a FIT in primary care.3 6 When NICE set the threshold at ≥10 µg Hb/g faeces there was a paucity of data available on FIT in symptomatic patients tested in primary care prior to referral for colonic investigation.7 Since 2017, numerous studies have been published to inform the choice of FIT threshold in primary care,8–20 26–25 including some showing that FIT outperforms symptom-based referral criteria.24–26

Although the choice of FIT thresholds has an important role in the allocation of resources, the primary rationale for the selection of an fHb threshold is to ensure that a symptomatic population is offered interventions relative to absolute CRC risk. Symptomatic populations with an fHb below a defined threshold may theoretically not be indicative of increased CRC risk compared with an asymptomatic population, and therefore we sought to determine this risk according to fHb concentration.

Randomised controlled trials

There have been no RCTs comparing time to diagnosis, stage at diagnosis or longer-term CRC outcomes between patients with and patients without FIT as part of their diagnostic pathway, nor any trials comparing FIT based pathways using different FIT thresholds.

Systematic review of cohort studies of primary care patients

A systematic review summarised the diagnostic performance of FITs for CRC across a range of thresholds, including 69,536 symptomatic adults from primary care from 23 cohort studies published between May 2018 and November 2020.26 Using the reported limit of detection (LoD), which ranged from ≥2 µg Hb/g faeces to ≥7 µg Hb/g faeces, meta-analysis of 11 studies (n=41,388 patients) resulted in a pooled sensitivity of 93.4% (95% CI 88.0% to 96.4%) and specificity of 76.9% (95% CI 67.7% to 84.0%). At a threshold of ≥10 µg Hb/g faeces (15 studies; n=48,872), pooled sensitivity was lower at 87.0% (95% CI 81.0% to 91.6%) with higher specificity 84.4% (95% CI 79.4% to 88.3%). Meta-analysis of five studies (n=24,187) reporting at ≥20 µg Hb/g faeces resulted in a reduced sensitivity of 84.1% (95% CI 78.6% to 88.4%) and an increased specificity of 86.6% (95% CI 75.6% to 93.1%). At a threshold of ≥150 µg Hb/g faeces meta-analysis of six studies (n=34,691) resulted in a sensitivity of 64.1% (95% CI 57.8% to 69.9%) and a specificity of 95.0% (95% CI 91.2% to 97.2%).

Primary care cohorts with low prevalence of CRC

The underlying prevalence of CRC directly influences a test’s performance at a chosen threshold. Individual prospective and retrospective cohort studies reporting FIT performance in populations of patients tested in primary care report CRC prevalence ranging from 0.8% to 1.8%, highlighting the variation in symptomatic patient groups eligible for FIT across primary care settings.8 11 13 18 20 21 In a subgroup analysis of the review, at a threshold of ≥10 µg Hb/g faeces, sensitivity was lower at 86% (95% CI 78% to 93%) versus 89% (95% CI 82% to 96%) and specificity significantly higher at 87% (95% CI 82% to 92%) versus 81% (95% CI 74% to 88%) when eight studies with a combined prevalence <3% were compared with seven studies with a prevalence of ≥3%.26

Trade-offs between single FIT thresholds in primary care cohorts

Most cohort studies have reported the use of a single FIT threshold with some including statistical modelling to demonstrate the
trade-offs at different FIT thresholds in terms of the numbers needed to scope (NNS) to detect one cancer and the number of missed cancers (NMC) per 1000 patients tested.

The review reported that for a prevalence of 1% and 2%, the NNS was 20 and 10 using a threshold of ≥10 µg Hb/g faeces, and the NMC was 1 and 3 per 1000 patients, respectively. Increasing the threshold to ≥20 µg Hb/g faeces reduced the NNS to 12 and 6 and the NMC increased to 2 and 4 per 1000 patients tested. At 150 µg Hb/g faeces the NNS is reduced further to 7 and 4, and the NMC increased to 4 and 8 per 1000 patients tested.

Based on a large retrospective cohort of 9896 patients tested in English primary care in the context of the DG30 NICE guidelines (CRC prevalence 1.1%), the authors illustrated the NNS and NMC for the thresholds of ≥7, 10, 20, 50, 100, 120 and 150 µg Hb/g faeces. The corresponding proportion of positive tests were 11, 10, 7, 4, 3, 3 and 2%, the proportion of cancers detected 91, 91, 85, 74, 61, 57 and 54%, the NNS to detect one cancer was 11, 10, 8, 6, 5, 5 and 4, and the NMC per 1000 FITs was 1, 2, 3, 4, 5 and 5. Reducing the threshold of ≥10 to ≥2 µg Hb/g faeces resulted in an increase in the NNS to 21 and a reduction in the NMC to 4 per 1000 patients tested.

At a higher prevalence of 1.6%, a smaller Spanish retrospective cohort of 4543 symptomatic patients reported the NMC per 1000 patients tested to be 3.7 (2.2–6.3) at a threshold of ≥10 µg Hb/g faeces compared with 4.1 (2.5–6.6) using ≥20 µg Hb/g faeces, and the NNS was 13.8 (10.8–17.7) compared with 10.9 (8.5–14.0). The authors concluded that the use of ≥20 µg Hb/g faeces in preference to ≥10 µg Hb/g faeces could reduce referrals for colonoscopy without missing more than one CRC per 1000 patients tested.

Increasing the threshold favours specificity reducing the NNS to detect one CRC but increases the NMC per 1000 patients tested. The opposite occurs when the threshold is reduced. The FIT threshold used in clinical practice is likely to be chosen based on a balance of tolerance of missed cancers and the diagnostic resources available to urgently investigate CRC.

**Multiple thresholds for low-risk, intermediate-risk and high-risk populations**

Multiple FIT thresholds have been introduced in some clinical settings to provide a rule-out threshold, a rule-in threshold and an intermediate range where the investigation of population subgroups and/or active safety netting is advised.

The Nottingham rapid CRC diagnosis (RCCD) service triages adult patients of any age, except those with rectal bleeding and rectal mass, combining low, intermediate and high thresholds (prevalence 1.6% (227/13 361)). The RCCD service considers ≥100 µg Hb/g faeces the ‘high risk’ positive contacting these patients directly for rapid investigation. Patients with a FIT result <4 µg Hb/g faeces, and with a FIT result of 4–10 µg Hb/g faeces but normal blood tests are considered ‘negative’. Patients with a FIT of 4–10 µg Hb/g faeces with anaemia, low ferritin or thrombocytosis, or with a FIT ≥10 µg Hb/g faeces are considered ‘positive’ and investigated urgently via the 2WW. The cancer detection rate was 0.1% for <4 µg Hb/g faeces, 0.6% between 4 and 9.9 µg Hb/g faeces, 3.3% for 10–99.9 µg Hb/g faeces and 20.7% for ≥100 µg Hb/g faeces.

Cohort data from Tayside on FIT use in primary care patients with unselected GI symptoms was modelled to show that ‘reassurance thresholds’ of <2, 7, 10 and 20 µg Hb/g faeces were associated with CRC risks of 0.1, 0.3, 0.3 and 0.4%. Intermediate-risk populations were created to highlight those with a cancer risk below the 3% risk used by NICE to trigger urgent colorectal investigation. For example, an intermediate population defined by 10–99 µg Hb/g faeces had a risk of 2.7%, leaving a higher risk population ≥100 µg Hb/g faeces with a risk of 14.5%. An intermediate range of 10–149 µg Hb/g faeces resulted in an intermediate population risk of 3.2%, meaning all patients ≥10 µg Hb/g faeces would qualify for urgent investigation. The cancer risk for the intermediate 7–199 µg Hb/g faeces population was 2.8% with a risk of 17.2% in the ≥200 µg Hb/g faeces group. However, there was no intermediate population ≥20 µg Hb/g faeces with a risk below 3%. The patients with FIT ≥20 µg Hb/g faeces comprised 16.8% of the population tested compared with 21.9% for ≥10 µg Hb/g faeces and 25.4% for ≥7 µg Hb/g faeces.

**Individualised FIT thresholds**

All dichotomous FIT thresholds, from the LoD upwards, identify a population with CRC risk ≥3% as recommended by NICE for urgent referral. For example, the PPV for ≥2 µg Hb/g faeces was 4.7% (4.0% to 5.5%) in the Oxfordshire Primary Care cohort, rising to 8.4% (7.1% to 9.9%) using ≥10 µg Hb/g faeces. An analysis from the Southwest of England discussed moving from population risk to individual risk. The cancer risk in the ≥10 µg Hb/g faeces population was 7%, in line with larger data sets from low prevalence primary care populations. Although cautious about the uncertainty in their estimate, they calculated that the individualised risk was 3% at the threshold ≥37 Hb/g faeces (95%CI 26 to 50) suggesting safety netting may be warranted between 10 and 36 Hb/g faeces.

**SHOULD FIT BE USED IN PRIMARY OR SECONDARY CARE?**

We recommend that FIT should be used by primary care clinicians to prioritise patients with clinical features of CRC for referral for urgent investigation.

**GRADE of evidence: very low; Strength of recommendation: Strong.**

There are no controlled studies or economic evaluations that compare the effectiveness of pathways using FIT in primary care with pathways using FIT in secondary care.

**FIT in primary care: pre-referral**

Four large retrospective cohort studies, in which not all individuals are investigated but have been followed-up, have described FIT usage for symptomatic patients in primary care: Northern Spain (n = 38 675), Oxford (n = 16 604), Tayside (n = 5372) and Nottingham (n = 24 855) where palpable rectal mass and bleeding were excluded.

In these large low prevalence cohorts, the cancer diagnosis rate at follow-up after reassurance without investigation based on very low Hb levels in primary care was 0.3% or less, regardless of variations in the platforms and cut-offs used. A further Danish study of 3462 patients evaluated FIT in patients without ‘alarm symptoms’ similarly found that the risk of CRC was <0.1% in those with Hb <10 µg Hb/g faeces.

Evidence from these studies also demonstrates that primary care clinicians will still refer patients to secondary care where clinical concern persists via appropriate or alternate pathways. Unpublished data from Nottingham suggests one in three patients are seen in an alternate pathway after ‘negative’ FIT. One in seven FIT below the threshold patients were investigated in Denmark. A study from Southwest England, reported that GPs made a referral within 3 months for 1 in 10 negative FITs.
detecting more than half of the FIT below the threshold CRCs (5 of 8).

Some, not all, of these populations were included in a pooled analysis of 15 studies including 48,872 patients,25 yielding a sensitivity for CRC of 87.2% (95% CI 81.0% to 91.6%) when using a threshold of ≥10 µg Hb/g faeces. A threshold of ≥20 µg Hb/g faeces missed less than one additional CRC per 1000 patients (from a population of five studies; n=24,187, with CRC prevalence 2%).

FIT in secondary care: post-referral
Large UK cohort studies of patients preselected for referral by GPs describe the performance of FIT in populations with the majority fulfilling ‘high-risk’ NG12 criteria all receiving colonic investigation.

The NICE FIT study,28 a multicentre double-blinded study of 9822 patients undergoing colonoscopy, demonstrated that the risk of bowel cancer was around 0.2% in those with undetectable levels of fHb and 0.4% in those with fHb <10 µg Hb/g faeces. The Fast Track FIT study29 evaluated 5040 patients undergoing colonoscopy, CTC or colorectal telephone assessment pathway showed the risk of bowel cancer was 0.4% in undetectable fHb and 0.5% in those with fHb <10 µg Hb/g faeces. The quantitative FIT (qFIT) study30 reported on 3596 patients who underwent colonoscopy or CTC and the risk of CRC was 0.4% in those with undetectable fHb and 0.5% in those with fHb <10 µg Hb/g faeces. FIT kits were provided to patients in both primary care and hospital settings. A Scottish study of 4841 referred patients reported the risk of CRC when fHb <10 µg Hb/g faeces was 0.6%.31

Overall, this evidence suggests that FIT has an acceptable miss rate whether used in primary care prior to referral or in secondary care following referral.

FIT in primary care vs no FIT
A service evaluation from Nottingham including 1668 patients provides a small real-life uncontrolled comparison of an NHS Trust that adopted FIT for symptoms (excluding rectal bleeding and palpable rectal mass) covering half of the regions urgent 2WW CRC referrals and a private provider for the remainder of the referred population.9 FIT rollout increased 2WW referrals and the proportion of new CRC diagnoses made on 2WW pathways. Emerging differences were noted in the cost of investigations required to detect each CRC and the time to diagnosis favouring the pathway using FIT.32 Other regions of the East Midlands have restricted FIT usage to those over 60 years and demonstrated demand reduction33 similar to the experience in Scotland where the introduction of FIT achieved a 15% reduction in urgent referrals from primary care.34

In Northern Spain35 the clinical outcomes of 279 patients with symptomatic CRC diagnosed after a ‘positive’ FIT in primary care were compared with 1210 patients with symptomatic CRC without a primary care FIT. A higher proportion of Stage I and II cancers (51.3% vs 45.5%) and improved 3-year survival were found in the FIT group. The Nottingham group9 reported a pre-pandemic shift towards diagnosis at earlier stage. Juul et al36 report 66.7% of cancers diagnosed at Stages I and II when evaluating FIT in symptomatic patients without ‘alarm symptoms’. Bailey et al37 found 33% were detected at early stage when using FIT in those that satisfied DG30 criteria specifically. Turvill et al29 described a higher proportion of Stage I and II CRC (52.7%) in their referred Fast Track study population with low fHb <18 µg Hb/g faeces although numbers are very small. Other studies have shown obstructing tumours and higher stage CRC are also common in FIT ‘negative’ CRC.

It is not possible to conclude that introducing FIT in primary care improves longer-term outcomes, but evidence is emerging that FIT testing in primary care could have this impact.

FIT only in secondary care
During the pandemic FIT was adopted widely across the UK given fears that endoscopy and CTC were aerosol generating procedures that increased risk of viral transmission.38 High fHb thresholds (100 µg Hb/g faeces in England and Wales, 400 µg Hb/g faeces in Scotland) were recommended with a pragmatic acceptance that some diagnoses would be missed, and a number of reports have described this.

High fHb could be used to identify referred patients for urgent/2WW/direct to colonoscopy pathways, with lower fHb directed to routine pathways. FIT could be used to ‘upgrade’ patients referred routinely and fHb might be used to determine the choice between colonoscopy and CTC, or colonoscopy and flexible sigmoidoscopy. An fHb could also become a component of all informed consent conversations for invasive colonic investigation. The current statute law underlying the 2WW system and the timed nature of pathways inhibits this individualised approach. ‘Local agreement’ is key to best practice in implementation.

It has been suggested that access to FIT should be restricted for use in secondary care but studies describing a favourable stage shift associated with FIT use suggest that a secondary care only approach may miss the opportunity to increase the proportion of cancers detected by identifying higher risk patients before referral.

ADVICE FOR CLINICIANS WHERE PATIENTS HAVE NOT RETURNED A FIT TEST
We suggest that clinicians should follow-up patients with no FIT result to encourage them to return a sample or, where the kit has been lost or inadequately submitted, offer a further test. GRADE of evidence: very low; Strength of recommendation: weak.

We suggest that patients who decline to return a FIT test should be counselled that evaluation of their symptoms is incomplete, and be encouraged to complete their test. GRADE of evidence: very low; Strength of recommendation: weak.

We suggest that where no FIT result can be obtained, clinicians should use existing national and local guidelines to assess risk of CRC. GRADE of evidence: very low; Strength of recommendation: weak.

There is very limited evidence on how to manage patients who do not return and/or refuse to undertake a FIT test, however we have sought to develop relevant advice for clinicians.

There is limited survey evidence that most patients find FIT testing acceptable but that people from ethnic minorities may be less likely to return kits possibly due to concerns about hygiene. There are a few studies suggesting possible interventions that may improve rates of return of FIT kits. In the absence of abdominal or rectal mass or ulceration, FIT is the best discriminator of a patient’s risk and need for referral to investigate possible CRC.

Importance of FIT testing
Numerous studies and reviews have found strong associations between a FIT positive test and risk of CRC. Studies that
compared the strength of this association with other risk factors have shown that a positive FIT test is usually more predictive of risk than demographic, clinical and laboratory criteria for referral to investigate possible CRC. 28 34–36

**Improving return of tests**

A small number of articles were identified that provide insights into strategies for maximising return of FIT kits in patients who have been asked to have the test.

Coronado et al undertook surveys on worded versus non-worded instructions for performing FIT testing in an area with a population having high rates of non-English speakers. They reported preferences for non-worded instructions from both patients and professionals.37

A US study of FIT used for screening found that patients asked to complete a two-sample test were less likely to return than those completing a one-sample test; this difference was statistically significant though not numerically large (39.6% vs 43.3%).38

Haghighat et al described an initiative to encourage patients to submit their FIT kit as part of colorectal screening as soon as possible after this was offered, ideally before leaving the clinic. They reported significant improvements over the 6-month period of study (27.6% vs 20.6%, p<0.001).39

A survey of patients in the NICE FIT study examined 1151 questionnaires representing 30.6% of those mailed out (1151/3760), with lower percentage returns from London than outside London patients (17% and 43%, respectively). Most patients found FIT collection straightforward (90.2%), not unhygienic (76.3%) and preferable to colonoscopy (78.1%). People aged 40–64 years were less likely to prefer FIT to colonoscopy than older age groups. Patients from ethnic minority backgrounds were less likely to have found the test hygienic and to return the kit in a future test.40

**SAFETY NETTING**

We recommend that some patients with symptoms of suspected CRC may be managed in primary care if fHb <10µg Hb/g, and provided appropriate safety netting is in place

GRADE of evidence: very low; Strength of recommendation: strong.

We suggest that patients with an fHb <10µg Hb/g but with persistent and unexplained symptoms for whom the GP has ongoing clinical concern should be referred to secondary care for evaluation.

GRADE of evidence: very low; Strength of recommendation: Strong.

We recommend that safety netting protocols should incorporate advice and strategies for the diagnosis of CRC and extracolonic cancer, as well as other serious gastrointestinal conditions.

GRADE of evidence: very low; Strength of recommendation: weak.

Most patients with symptoms and signs suggestive of CRC may be managed in primary care if the FIT level is low or undetectable. The risk of CRC in patients with fHb <10 µg/g of faeces approximately equates to the risk of severe complications from colonoscopy, or to the CRC risk in asymptomatic subjects. GPs may consider alternative causes for abdominal symptoms if FIT testing is below the threshold for referral, given the absolute risk of CRC is low in this situation. Where the primary care clinician has ongoing concerns about a serious cause, they should consider non-GI as well as GI conditions, including cancers not located in the colon or rectum. There appears to be a limited role for other tests such as a full blood count (FBC), combined with persistent symptoms and/or clinical acumen, to determine which patients with a negative FIT result may be considered at increased risk of CRC, and further work may be required to identify robust safety netting mechanisms which could be employed. However there will always remain an important role for clinical acumen and personalisation of care for patients who may be managed in a myriad of different ways for very different symptoms, and who may refer on alternate pathways, or managed in primary care accordingly (figure 1).

**Safety netting**

Safety netting has come to be regarded as ‘best practice’ in relation to cancer diagnosis, especially in non-specialist settings.41 42 Its aim is to ensure patients do not drop through the healthcare net but are monitored until symptoms are explained, defined as a consultation technique to communicate uncertainty, provide patient information on red-flag symptoms and plan for future appointments to ensure timely re-assessment of a patient’s condition.43 NICE refer to safety netting as ‘the provision of support for patients in whom the clinician has some uncertainty as to whether the patient has a self-limiting illness and is concerned that their condition may deteriorate’.44 However, safety netting may also comprise administrative activities such as test result reconciliation and the follow-up of referrals.44 45 A key role for safety netting in FIT based pathways is the monitoring of FIT negative patients to ensure timely referral or investigation of those referred.

**Absolute risk of missed cancers**

In a recent meta-analysis pooling data from 35 925 patients from nine UK studies from primary and secondary care including a range of FIT thresholds from ≥2 to ≥19 µg Hb/g faeces and within the NICE NG12 context, the pooled percentage of missed CRCs due to a FIT below the selected threshold was 8.7% (95% CI 5.1% to 12.2%) equating to an NMC of 2.1 per 1000 patients tested.46 Pin-Vietto et al reported the NMC as 1 per 1000 patients tested at a CRC prevalence of 1% and 3 in 1000 at a prevalence of 2%, using a threshold of ≥10 µg Hb/g faeces, noting that the prevalence of CRC in primary care based studies ranged from 0.8% to 1.8%.26 Risks of 1 and 3 per 1000 patients equate to absolute cancer risks of 0.1% and 0.3%, respectively, both significantly below the current NICE threshold of ≥3% cancer risk used to warrant urgent cancer investigation.47 An effective safety netting strategy could identify FIT negative patients with an increased risk of cancer who may warrant further investigation.

**Meta-analyses and randomised controlled trials**

We found no meta-analyses or RCT of safety netting strategies to ensure CRCs are diagnosed in patients with a negative FIT, in primary or secondary care settings. A step-wedged cluster RCT is currently underway to test the effectiveness of an electronic safety netting toolkit embedded into major primary care clinical systems to facilitate patient follow-up in terms of the time and route to diagnosis.48

**Observational studies**

There were no observational studies evaluating safety netting strategies to promote re-consultation and onward referral among people with ongoing symptoms despite a negative FIT result.

Recent research has emphasised the importance of clinician ‘gut feeling’ in the diagnosis of cancer, conceptualised as the
rapid summing up of multiple verbal and non-verbal patient cues.\textsuperscript{49} FIT pathways should allow clinicians to refer or investigate FIT negative patients if there are ongoing clinical concerns. In a cohort study from the Southwest of England, GPs still requested urgent investigation for five of the eight FIT negative cancers ‘probably because continuing symptoms allowed the GP to ‘override’ the negative test’.\textsuperscript{11}

A common suggestion in the literature and guidelines was re-consultation within 4–6 weeks for patients with ongoing symptoms and a FIT below threshold. G27 guidance (2005), later replaced by NG12, recommended urgent referral for patients with symptoms persistent for 6 weeks. Underpinning evidence to inform the timing of any safety netting action is lacking including: what might be considered the ‘normal’ duration of a benign symptom, the time taken for progression of high-risk adenomas to cancer, or the interval of stage progression.

Many cohort studies have documented the clinical presentation of patients later diagnosed with FIT negative CRC, suggesting that these characteristics could be prioritised for referral or included in the communication of safety netting advice to patients.\textsuperscript{16,27,50,51} However, there is marked variation in the characteristics of FIT negative cancers between these studies and so relying on these characteristics to inform a safety netting strategy could be falsely reassuring.

**Modelling studies**

Modelling studies to date have not demonstrated the benefit of combining FIT with other clinical features and blood test results to enhance sensitivity by reducing false negative FITs. A comparison of FIT at ≥10 μg Hb/g faeces alone, with the FAST score (combining FIT age and sex), and ColonFlag (a machine learning algorithm using age, sex and FBC indices to derive a risk score), showed that FIT and ColonFlag missed a different 18% of CRCs, respectively, and FAST score missed 27.3%.\textsuperscript{13} Combining simple blood tests with FIT at best matches the sensitivity of FIT alone in patients tested in primary care, whether as pairs of results or within multivariable model.\textsuperscript{20}

**IS A REPEAT/SECOND FIT USEFUL AND DOES IT ENHANCE DIAGNOSTIC ACCURACY?**

Studies suggest repeat FIT testing may enhance sensitivity, but lower specificity, and this depends on whether the second test is used to identify people to be investigated/referred after the first test is negative (increased sensitivity and decreased specificity) or to identify people who may not need referral unless both tests are positive (decreased sensitivity and increased specificity).

Studies have examined cohorts identified for investigation (or already investigated/diagnosed) rather than prospectively using FIT to guide referral in ‘real world’ situations. Although the populations under study have varied considerably (symptomatic vs screening; high vs low risk) the findings of sensitivity and specificity have been relatively consistent. No studies were found that examined the optimal period for undertaking a repeat/second FIT test. Where it was clear in the methods, most studies instructed repeat FITs to be sampled from consecutive stools. In conclusion, although there is currently insufficient evidence to recommend use of repeat/second FIT to guide referrals in routine practice, further data are required to clarify the role of this approach (online supplemental file 2).

There have been no randomised controlled trials or systematic reviews comparing diagnostic yield, time to diagnosis, stage at diagnosis or longer-term CRC outcomes between patients who have one and those having repeat/second FIT tests. Mosen et al\textsuperscript{18} conducted an RCT of 3121 participants comparing uptake of a two sample regimen (1562) with one sample FIT (1559). Participants were given the same instructions with the 2-FIT group required to do the test twice. The FIT was requested in the context of bowel cancer screening. No significant difference in the baseline characteristics of each group. The FIT kits were posted to the participants and returned by mail. A total of 43.3% of the 1-FIT group were compared with 39.6% (p=0.012) of the 2-FIT group. In a large systematic review summarising the diagnostic performance of FITs for CRC including 69 536 symptomatic adults from primary care including 23 cohort studies, where studies included findings from patients tested more than once, only the first FIT result was analysed.\textsuperscript{1} Therefore the predominant evidence source is observational.

Turvill et al\textsuperscript{52} undertook a prospective, blinded observational study of associations between FIT results from two samples in all patients referred to York Hospital with suspected CRC within the urgent (2WW) pathway from February 2016 to March 2017. The FIT samples were provided by the patients between the hospital clinical appointment and investigations in secondary care. For patients with a single positive FIT, a threshold of ≥10 μg/g, was associated with sensitivity of 84.6% and specificity of 88.7%. For patients with two positive FIT tests, sensitivity was 91.7% and specificity of 85.1%. The paper did not examine and compare sensitivity and specificity in patients who were only offered one FIT test. Nor did it report on the negative predictive value (NPV) of two versus one FIT <10μg/g.

Hunt et al\textsuperscript{54} examined the association between CRC diagnosis and FIT results in patients who had two FIT following a referral to a specialist service from 2017 to 2021.\textsuperscript{12} Patients had been referred under different clinical pathways at different times, specifically at times under a low-risk versus high-risk pathway. The patients were asked to return two FIT kits from different stools before clinical assessment in secondary care. They found that if patients had been referred based on FIT result, sensitivity for CRC would have been 97.8% and 91.5%, specificity 66.2% and 81.6% and PPV 3.1% and 5.2% with one or two FIT positive test results (>10 μg/g), respectively. Two tests were returned by 96.1% of the study population indicating the data were representative of the study population. Those studies were initially those referred with low-risk symptoms and later those with high-risk symptoms. Missed CRC detection with two FIT <10 was found in 7/73 (9.6%). All the patients with ‘missed’ cancers had anaemia and one had an obstructing tumour.

This study provides evidence that a requirement to test positive (FIT >10μg/g) twice rather than once before decision to refer, or to investigate post-referral, may reduce the numbers of people referred or investigated, respectively, though at the cost of missing a proportion of CRC cases. The experience may not directly test real-world practice however where those with one or even two negative tests may still be referred and/or investigated if there remain clinical concerns.

Mattar et al\textsuperscript{56} studied 289 patients who underwent colonoscopy who had been entered into either a one-sample or two-sample FIT protocol.\textsuperscript{13} It is not clear from the description if patients selected were from a symptomatic or screening population. Among them 172 had one-sample FIT; for these positive and negative rates were not reported but colonoscopy outcome findings were reported in 99 cases and 117 people received the 2-sample FIT and of these 94 (80.3%) patients had both FIT below the threshold, 13 (11.1%) had both FIT positive and 10 (8.5%) had only one FIT positive (≥10μg/g). For the one-sample FIT group, positive FIT had a sensitivity and specificity of 83.3% and 86.9%, respectively. For the two-sample group,
those who had at least one sample positive had sensitivity and specificity of 75% and 92.9%, respectively. Separate figures for those who had two versus one sample positive in the two-sample group were not provided in the text.

Observational studies comparing the use of one and two FITs in bowel cancer screening reveal increased PPV for CRC and high-risk polyps. Moosavi et al. reviewed 17,031 participants in the British Columbia screening programme. The PPV following two positive FITs (at a cut-off of 20 μg/g) was 8% versus 1% for one positive FIT. For high-risk polyps the PPVs were 40% and 20%, respectively. CRC and high-risk polyps were missed with one FIT specimen. Polyps amounting to 12.1% of cancers and 23.4% were identified in patients where the first FIT was negative and the second positive.

Lim et al. retrospective study of 1,672 participants from Singapore’s bowel cancer screening programme found one FIT cohort had significantly less cases of CRC and polyps found than in the two FIT group. Both these studies asked patients to sample stools on consecutive days. As investigation of the bowel in these studies is only triggered by a FIT result over the threshold, the negative predictive of two FIT compared with one is not known.

In a population-based case-control study, Kim et al. examined the associations between previous colonoscopy and FIT testing, and the risk of future CRC diagnosis by comparing data from 61,221 patients with newly diagnosed CRC (case group) and 306,099 individuals without CRC (control group). Data on testing and diagnosis were from claims data from the Korean National Health Insurance System. They found that previous FIT testing was associated with lower OR for CRC, but where patients had records of >1 previous FIT ORs successively increased. FIT testing may have largely reflected screening so may not be applicable to use in symptomatic patients. Where symptoms based, repeat testing may reflect the presence or persistence of concerning symptoms rather than the utility of repeat testing.

Maeda et al. evaluated the impact of using one versus two FIT tests (using ≥10 μg/g threshold) to guide specialist investigation by modelling in a COVID-19 adapted pathway. The study also examined the impact of CT mini-prep. Values for FIT sensitivity and specificity used in the analysis were derived from audit data, South East Scotland Cancer Network data, literature and, if missing, assumptions on reasonable (best-versus-worst-case) scenarios were made by expert opinion. Sensitivity (84%) and specificity (74%) figures were broadly in keeping with those reported by cohort studies. The modelling estimated that investigating all patients with any positive FIT result out of two would reduce the risk of missing a CRC from 20.2% to 15.3%, identifying 13.3 versus 10 patients per 1000 patients referred on a non-FIT pathway, while increasing the numbers investigated from 287 to 359 per 1000 patients on the referral pathway. This is compared with fewer than 5% missed CRC (3 per 1000) on the pre-COVID pathway.

The COLONFIT study developed a scoring system to prioritise fast-track colonoscopy. They obtained three FIT samples from 1,495 patients with symptoms (1,058 met NICE NG12 guidelines) diagnosing 116 CRC. 6/116 (5%) had only 1/3 FIT; 11, 3 CRC patients (2.6%) had negative FIT <4 and 2 patients were >4 μg/g and <11 μg/g.

**DIAGNOSTIC ACCURACY OF FIT FOR CRC IN PEOPLE WITH SUSPECTED CRC SIGNS OR SYMPTOMS**

FIT is a triage tool to identify those patients with symptoms of suspected CRC who should undergo further colorectal investigation.

| Table 1 | Number needed to scope (NNS) to detect one cancer and number of missed cancers (NMC) per 1000 faecal immunochemical tests (FITs) at various thresholds of FIT |
|------------------|------------------|-----------------|------------------|-------------------|
| Threshold (μg Hb/g) | Positive FITs n (%) | Negative FITs n (%) | Cancers detected n (%) | NNS to detect one cancer | NMC per 1000 FITs |
| ≥7 | 111 (11) | 889 (89) | 10 (91) | 11 | 1 |
| ＞10 | 96 (10) | 904 (90) | 10 (91) | 10 | 1 |
| ≥20 | 71 (7) | 929 (93) | 9 (85) | 8 | 2 |
| ≥50 | 44 (4) | 956 (96) | 8 (74) | 6 | 3 |
| ≥100 | 30 (3) | 970 (97) | 7 (61) | 5 | 4 |
| ≥120 | 28 (3) | 972 (97) | 6 (57) | 5 | 5 |
| ≥150 | 25 (2) | 975 (98) | 6 (54) | 4 | 5 |

**GRADE of evidence:** low; **Strength of recommendation:** strong.

We suggest that FIT be used for people with iron deficiency anaemia within primary care to inform urgency of referral.

**GRADE of evidence:** low; **Strength of recommendation:** weak.

We suggest referral of patients with persistent/recurrent anorectal bleeding for flexible sigmoidoscopy if fHb <10μg Hb/g.

**GRADE of evidence:** very low; **Strength of recommendation:** weak.

In summary, a meta-analysis informing these guidelines and four prior meta-analyses of over 48,000 patients, which include the largest three studies from NICE FIT, qFIT and the York groups on FIT diagnostic accuracy reported a FIT sensitivity for CRC in symptomatic patients to be greater than 87% at a threshold of 10 μg/g (table 1). Therefore, based on the studies reviewed in this section, FIT should be considered in patients presenting with lower GI symptoms irrespective of their nature, to support referral or triage to appropriate investigations if there are concerns about a cancer diagnosis.

**Introduction**

Considerable emerging evidence has recently been published on the diagnostic accuracy of FIT in symptomatic patients using a range of thresholds. The most common reported fHb triage threshold is 10 μg/g. Many of these reports assessed all patients presenting to their primary care practitioners with any bowel symptoms suspicious of CRC, but others categorised symptoms into low and high risk symptoms as defined by NICE. Only a few studies examined the diagnostic accuracy of FIT for individual symptoms.

In this section we focus on those studies where all symptomatic patients that received FIT were linked with a reference standard investigation with full evaluation of the colon and rectum to exclude CRC with either colonoscopy or CTC. Where these evaluations were not performed, studies with other reference standards such as flexible sigmoidoscopy or CT scan and those with clinical follow-up/record linkage of a minimum of 3 months were examined. Where this was not available then studies with lesser follow-up and other linked investigations such flexible sigmoidoscopy or CT scan were included. In assessing the diagnostic accuracy of FIT we considered its accuracy in symptomatic patients in general, and also in ‘high-risk’ and ‘low-risk’ symptoms.

**Early studies (table 2)**

One of the first studies to examine the role of FIT in symptomatic patients was published in 2011 by a Dutch group who
Conducted a clinical study involving five centres testing FIT in a mixed cohort of symptomatic and asymptomatic patients who were scheduled for colonoscopy. A total of 2145 patients undifferentiated by colonoscopy indication were included which were then divided irrespective of their symptoms or lack of, into high and low risk groups for CRC. The overall sensitivity and specificity of FIT for CRC was 92.4% (95% CI 84.2% to 97.2%) and 86.4% (95% CI 84.8% to 87.9%). In 2013, the Tayside group in 2013 examined 280 participants who had both FIT and endoscopy results. The sensitivity of FIT was also higher than NICE and SIGN criteria (77.4%, 12.8%; p<0.001). In 2016 the Tayside group reported on 484 patients who had FIT and colonoscopy and reported that all 11 cancers were detected at a cut-off of 10 µg/g (100% sensitivity).

Later studies (table 3)

In 2017, NICE produced its DG30 guidelines which recommended the use of FIT in low-risk symptoms defined as ‘patients without rectal bleeding who have unexplained symptoms but do not meet the criteria for a suspected cancer referral’. The latter refers to high risk symptoms as defined in the updated NICE NG12 guidance in 2017 where use of FIT was not recommended. The DG30 guidance was based on a health technology assessment, commissioned by The National Institute for Health Research to produce a diagnostic accuracy report on FIT to triage symptomatic patients at low risk of CRC presenting in primary care. The report looked at 10 studies but summarised evidence from 5 studies only that reported on FIT as a rule-out test for CRC with a cut-off of 10 µg/g. Data were taken from one study (507 patients) for the HM-JACKarc analytical system, and four studies (4091 patients) for the OC-Sensor analyser. The summary estimate of sensitivity for the HM-JACKarc was 100% (95% CI 71.3% to 100%) and for the OC-Sensor was 92.1% (95% CI 86.9% to 95.3%). The corresponding specificity was 76.6% and 85.8%.

Table 2  FIT diagnostic accuracy for CRC at a cut-off of 10 µg/g in the early diagnostic accuracy studies prior to 2017

<table>
<thead>
<tr>
<th>Region</th>
<th>Design</th>
<th>n</th>
<th>Analyser</th>
<th>Reference standard</th>
<th>CRC sensitivity @ 10 µg/g</th>
<th>CRC specificity @ 10 µg/g</th>
<th>CRC PPV @ 10 µg/g</th>
<th>CRC NPV @ 10 µg/g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terharr sive Droste et al64</td>
<td>Holland</td>
<td>DTA</td>
<td>OC-Sensor</td>
<td>Colonoscopy</td>
<td>91.1% (84.2% to 95.6%)</td>
<td>87.0% (85.4% to 83.5%)</td>
<td>–</td>
<td>99.4% (98.9% to 99.7%)</td>
</tr>
<tr>
<td>McDonald et al64</td>
<td>UK (Scotland)</td>
<td>DTA</td>
<td>OC-Sensor</td>
<td>Colonoscopy</td>
<td>100% (54.1% to 100%)</td>
<td>93.8% (90.3% to 96.3%)</td>
<td>–</td>
<td>100% (98.5% to 100%)</td>
</tr>
<tr>
<td>Rodriguez-Alonso et al65</td>
<td>Spain</td>
<td>DTA</td>
<td>OC-Sensor</td>
<td>Colonoscopy</td>
<td>96.7% (82.8% to 99.9%)</td>
<td>79.9% (77.2% to 82.3%)</td>
<td>12.8% (9.1% to 17.9%)</td>
<td>99.9% (99.3% to 100%)</td>
</tr>
<tr>
<td>Mowat et al64</td>
<td>UK (Scotland)</td>
<td>DTA</td>
<td>OC-Sensor</td>
<td>Colonoscopy</td>
<td>89.3% (71.8% to 97.7%)</td>
<td>79.1% (75.9% to 82.0%)</td>
<td>14.2% (9.8% to 20.1%)</td>
<td>99.5% (98.5% to 99.8%)</td>
</tr>
<tr>
<td>Godber et al66</td>
<td>UK (Scotland)</td>
<td>DTA</td>
<td>HM-JACKarc</td>
<td>Colonoscopy</td>
<td>100% (71.5% to 100%)</td>
<td>76.6% (72.6% to 80.3%)</td>
<td>9% (5.1% to 15.4%)</td>
<td>100% (99.0% to 100%)</td>
</tr>
</tbody>
</table>

95% CIs given in brackets when reported.

CRC, colorectal cancer; DTA, diagnostic test accuracy study; FIT, faecal immunochemical testing; NPV, negative predictive value; PPV, positive predictive value.

Table 3  Systematic reviews of FIT diagnostic accuracy for CRC at a cut-off of 10 µg/g in symptomatic patients

<table>
<thead>
<tr>
<th>Design</th>
<th>n</th>
<th>Analyser</th>
<th>Reference standard</th>
<th>CRC sensitivity @ 10 µg/g</th>
<th>CRC specificity @ 10 µg/g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Westwood et al63</td>
<td>SR</td>
<td>4091</td>
<td>OC-Sensor</td>
<td>Various</td>
<td>92.1% (86.9% to 95.3%)</td>
</tr>
<tr>
<td>Pin Viento et al65</td>
<td>SR</td>
<td>4075</td>
<td>OC-Sensor</td>
<td>Various</td>
<td>94.1% (90.0% to 96.6%)</td>
</tr>
<tr>
<td>Stonestreet et al66</td>
<td>SR</td>
<td>4096</td>
<td>OC-Sensor</td>
<td>Various</td>
<td>93%* (88% to 99.6%)</td>
</tr>
<tr>
<td>Pin Viento et al65</td>
<td>SR</td>
<td>48872</td>
<td>OC-Sensor/HM-JACKarc</td>
<td>Various</td>
<td>87.2% (81.0% to 91.6%)</td>
</tr>
<tr>
<td>Saw et al67</td>
<td>SR</td>
<td>25500</td>
<td>OC-Sensor/HM-JACKarc/FOB Gold/QuickRead Go</td>
<td>Various</td>
<td>88.7% (85.2% to 91.4%)</td>
</tr>
<tr>
<td>Booth et al68</td>
<td>SR</td>
<td>35945</td>
<td>OC-Sensor/HM-JACKarc/FOB Gold/QuickRead Go</td>
<td>Colonoscopy and CT</td>
<td>91.0% (88.9% to 92.7%)</td>
</tr>
</tbody>
</table>

95% CIs given in brackets when reported.

*10–15 µg/g.

CRC, colorectal cancer; CTC, CT colonography; FIT, faecal immunochemical testing; SR, systematic review and meta-analysis.
Two years later, a meta-analysis by the Warwick group included 17 studies of which 9 were on symptomatic cohorts (6755 patients). Fifty studies used the OC-sensor (4883 patients), three used HM-JACKarc (1499 patients) and one used the Actim Facal Blood system. Five Studies (4603 patients; four OC-Sensor and one HM-JACKarc) examined FIT at a cut-off of 10 µg/g and the others looked at a range of cut-offs ranging from 7 to 50 µg/g. The overall pooled sensitivity and specificity for CRC were 0.90 (95% CI 0.87 to 0.92) and 0.87 (95% CI 0.83 to 0.90), respectively.

A subanalysis for studies that used OC sensor was performed. These studies examined multiple cut-off concentration values ranging from 10 to 40 µg/g. Analysis of the pooled sensitivity and specificity for an fHb cut-off range of 10–15 µg/g (4096 patients) showed sensitivity of 0.93 (95% CI 0.88 to 0.96) and specificity of 0.87 (95% CI 0.82 to 0.90). For the range between 20 and 40 µg/g, the pooled sensitivity and specificity of 0.87 (95% CI 0.84 to 0.90) and 0.89 (95% CI 0.84 to 0.92), respectively.

In the same year, a meta-analysis from Spain by Pin Vieito included 14 studies of which 7 were on symptomatic cohorts. Four studies (4035 patients) reported on FIT with a cut-off of 10 µg/g using the OC-Sensor; one of these studies was conducted in Scotland, and the others were conducted in Spain, or pooled data from studies from both countries. The pooled sensitivity for CRC was 94.1% (95% CI 90.0% to 96.6%) and specificity 66% (95% CI 47.1% to 80.1%).

A meta-analysis by the same group in 2021 included 23 studies (69,536 patients). The meta-analysis examined diagnostic accuracy of FIT at different thresholds. Fifteen studies (n = 48,872) reported on a cut-off of 10 µg/g. The pooled sensitivity and specificity for CRC was 87.2% (95% CI 81.0% to 91.6%) and 84.4% (95% CI 79.4% to 88.3%), respectively. Five studies (n = 24,187) reported on a cut-off of 20 µg/g and the pooled sensitivity and specificity were 84.1% (95% CI 78.6% to 88.4%) and 86.6% (95% CI 75.6% to 93.1%). Six studies (n = 34,691) assessed FIT as rule in test cut-off of >150 µg/g showing a sensitivity of 64.1% (95% CI 57.8% to 69.9%) and a specificity of 95.0% (95% CI 91.2% to 97.2%). The group concluded that FIT is the test of choice to evaluate patients with new-onset lower GI symptoms in primary healthcare.

The most recent meta-analysis published in December 2021 from New Zealand included 15 studies with a cohort of 28,832 patients, all of whom were prospectively recruited. Three studies (six HM-JACKarc, four OC-Sensor, one FOB Gold and two QuickRead Go analysers; n = 25,500) reported on FIT at a cut-off 10 µg/g. The summary sensitivity and specificity were 88.7% (95% CI 85.2% to 91.4%) and 80.5% (95% CI 75.3% to 84.8%), respectively. At the lower cut-off of the LoD (three studies; 15,160 patients), the summary sensitivity increased to 96.8% (95% CI 91.0% to 98.9%) but specificity reduced to 65.6% (95% CI 59.0% to 71.6%).

A meta-analysis performed to inform these guidelines included 31 studies up to March 2022 with a cohort of 79,566 patients. For ‘all symptoms’, ‘all analyser’ analysis and a reference standard of >90% receiving either colonoscopy or CTC (16 studies, n = 35,945), the summary sensitivity and specificity were 91.0% (95% CI 88.9% to 92.7%) and 75.2% (95% CI 69.6% to 80.1%), respectively, at a cut-off 10 µg/g.61

In the 2 years between the 2019 and 2021 meta-analyses, there has been an explosion in the number of studies reporting on FIT, with between 25,000 to 40,000 patients added to the latest meta-analyses. The Pin Vieito and Saw meta-analyses included two of the three largest diagnostic accuracy, multicentre studies that were conducted in England. The qFIT study (UCLH Cancer Collaborative) included 3596 patients with high-risk symptoms and reported a sensitivity of FIT for CRC at 83.3% (95% CI 75.6% to 91.0%) at cut-off of 10 µg/g using the OC-Sensor. The NICE FIT study included 9822 patients with high and low risk symptoms and at the same cut-off, reported a sensitivity for CRC of 90.9% (95% CI 87.2% to 93.8%), using the HM-JACKarc analyser. The third largest research study published in 2021 by the York group included 5040 patients with high-risk symptoms and was included in a meta-analysis to inform these guidelines.61 62 63 The sensitivity and specificity of FIT for CRC at 10 µg/g using the HM-JACKarc analyser was 87.4% (95% CI 81.0% to 92.3%) and 80.9% (95% CI 79.7% to 81.9%), respectively. The group considered an optimal threshold between sensitivity and specificity and calculated this at 19 µg/g with a sensitivity of 85.4% (95% CI 78.8% to 90.6%) and specificity of 85.2% (95% CI 84.1% to 86.2%).

High and low risk symptoms

Low risk

Up until NICE released its DG30 guidelines, most studies investigated FIT in patients that presented to clinicians with bowel symptoms that required investigations to rule out bowel cancer. Indeed, DG30 was based on studies that included patients with wide-ranging symptoms, not stratified by high or low-risk symptoms in accordance with NICE criteria.

Since then, many studies began to report on low-risk and high-risk symptoms. There are a small number of studies investigating low risk symptoms presenting to primary care. The largest studies included three service evaluations11 13 19 and one diagnostic accuracy study.72 The service evaluations are not true diagnostic accuracy but are pragmatic studies reflective of what happens in real-life practice in that not all patients receiving FIT undergo investigations and those investigated may not necessarily receive full colonic imaging such as in elderly and/or unfit patients who may have a CT scan or flexible sigmoidoscopy instead.

Juul in 2018 investigated FIT at a cut-off of 10 µg/g in patients presenting with non-alarm symptoms in general practice in a Central Denmark Region.11 In total 3462 patients had FIT and of these, 540 (15.6%) were positive. Of these, 416 patients (77%) underwent diagnostic investigation within 3 months and 51 cancers (PPV: 9.4% (95% CI: 7.0% to 11.9%)) were found. Of the 2922 patients with FIT below 10 µg/g only 418 (14.3%) underwent a diagnostic investigation during the same period and three cancers were found.

The same year Nicholson et al reported on 238 patients with low-risk symptoms in Oxfordshire who had both faecal occult blood test and FIT and were followed-up for up to 21 months.15 The sensitivity and specificity of FIT at 10 µg/g were 85.7% and 89.2%, respectively. The PPV was 19.4% and NPV was 99.5%.

Bailey in 2021 reported on FIT at a cut-off of 10 µg/g in patients presenting in primary care in the Southwest of England with low-risk symptoms.11 Of these, 618 (15.9%) patients tested positive and were referred for investigations within 12 months and 43 were diagnosed with CRC (PPV 7.0% (95% CI 5.1% to 9.3%)). Of 3272 with FIT <10 µg/g, 324 (9.9%) were referred and in these 5 had CRC within 12 months. Of those 2948 patients who were not referred within 12 months, 3 had cancers. NPV was 99.8% (CI 99.5% to 99.9%). Sensitivity was 84.3% (95% CI 71.4% to 93.0%) and specificity was 85.0% (95% CI 83.8% to 86.1%).

The NICE FIT study reported on 1994 patients (20.3%) of the 9822 patients studied who had low-risk symptoms and 634 (6.5%) had other symptoms warranting urgent referral.12 The
The use of FIT in rectal bleeding was investigated recently in 462 patients in NHS Tayside, Scotland. The positivity rate was 63.3% at a cut-off of 10 µg/g. The prevalence of cancer was 8.5% (25/293) with an fHb >10 µg/g compared with 0.6% (1/168) when fHb <10 µg/g. The sensitivity and specificity for CRC was calculated as 96.2% (95% CI 80.4% to 99.9%) and 38.3% (95% CI 33.7% to 43.0%), respectively. The one CRC in the cohort with an fHb <10 µg/g was in the descending colon and would have been detected by flexible sigmoidoscopy. Indeed, flexible sigmoidoscopy detected the majority of serious bowel disease (SBD) pathology (CRC, inflammatory bowel disease and advanced adenomas) except for four advanced adenomas (10 out of 14 SBD out of 168 patients). The authors concluded that patients with an fHb <10 µg/g and persistent rectal bleeding, can be safely investigated with flexible sigmoidoscopy.

Högborg et al in 2020 reported on qualitative (using three FIT samples) rather than quantitative FIT in 606 patients with rectal bleeding. The positivity rate was 42% with a sensitivity of 96.2% and NPV of 99.7% for CRC.

The NICE FIT study investigated the diagnostic accuracy of FIT in 3143 patients with rectal bleeding either alone or in combination with other symptoms compared with 6679 patients with non-rectal bleeding symptoms. The positivity rate of 26.9% at 10 µg/g was lower than the other two studies above but higher than the non-rectal bleeding group at 15.2%. The sensitivity and specificity of FIT in the rectal bleeding group for CRC was 96.6% (95% CI 92.2% to 98.9) and 76.6% (95% CI 75.0% to 78.1%). The PPV was 16.8% (95% CI 15.9% to 17.9%) and NPV 99.8% (95% CI 99.5% to 99.9%). In the bleeding cohort, there were five CRCs with an fHb <10 µg/g, four of which would have been detected with flexible sigmoidoscopy. The group concluded that the use of flexible sigmoidoscopy in patients with rectal bleeding and an fHb <10 µg/g, would reduce the risk of CRC to 0.03%.

The Booth et al meta-analysis informing the guidelines identified three studies (n=3665) that reported specifically on rectal bleeding and used colonoscopy or CTC as a reference standard. The summary sensitivity at a cut-off of 10 µg/g was 96.6% (95% CI 92.8% to 98.8%) and specificity 71.7% (95% CI 70.2% to 73.2).

**Iron deficiency anaemia**

The Nottingham study, which used a postal FIT in both high and low risk groups described above as per NICE NG12 and DG30 (but excluding rectal bleeding), reported higher fHb levels in those with IDA at 4.8 (0.8–34.1) µg/g compared with 1.2 (0–6.4) µg/g in those without IDA. In this study 40 patients with CRC were identified and using a cut-off of 10 µg/g of fHb, CRC detection rate was 7.2 in those with IDA. In fact, the authors concluded that CRC detection was higher in those with IDA.

Cumin et al reported in their cohort of patients where FIT was used in primary care applying the NICE NG12 criteria that 7/48 patients (14.6%) had CRC below the cut-off of 10 µg/g, that is, fHb below threshold cancers. Of the seven fHb below threshold cancers, five had anaemia as well as change in bowel habits and of these, four had true IDA. It was also observed that these six CRCs were right-sided (caecal). The FIT sensitivity for CRC was 80.0% (95% CI 55.7% to 93.3%) in patients with IDA compared with 89.0% (95% CI 70.0% to 97.1%) in those with a combination of other symptoms.

Earlier studies did not report influence of IDA on diagnostic accuracy of FIT. More recent data in abstract form applying a definition of IDA of ferritin under 15, reported sensitivity of...
Table 4  FIT diagnostic accuracy for CRC at a cut-off of 10 µg/g in patients with high-risk symptoms in the largest and most relevant studies

<table>
<thead>
<tr>
<th>Region</th>
<th>Design</th>
<th>Symptoms</th>
<th>n</th>
<th>Analyser</th>
<th>Reference standard</th>
<th>CRC sensitivity @ 10 µg/g</th>
<th>CRC specificity @ 10 µg/g</th>
<th>CRC PPV @ 10 µg/g</th>
<th>CRC NPV @ 10 µg/g</th>
</tr>
</thead>
<tbody>
<tr>
<td>NICE FIT 2021</td>
<td>UK (England)</td>
<td>DTA</td>
<td>NG12</td>
<td>7194</td>
<td>HM-JACKarc</td>
<td>Colonoscopy</td>
<td>92.2% (88.2% to 95.2%)</td>
<td>88.3% (81.2% to 93.2%)</td>
<td>16.2% (13.4% to 18.2%)</td>
</tr>
<tr>
<td>Turvill et al 2021</td>
<td>UK (England)</td>
<td>DTA</td>
<td>All</td>
<td>5040</td>
<td>HM-JACKarc</td>
<td>Various</td>
<td>87.4% (81.0% to 92.3%)</td>
<td>80.9% (79.7% to 81.9%)</td>
<td>12.4% (10.4% to 14.5%)</td>
</tr>
<tr>
<td>qFIT 2021</td>
<td>UK (England)</td>
<td>DTA</td>
<td>NG12</td>
<td>3596</td>
<td>OC-Sensor</td>
<td>Various</td>
<td>83.3% (75.6% to 91.0%)</td>
<td>80.1% (78.9% to 81.4%)</td>
<td>9.7% (7.6% to 11.8%)</td>
</tr>
<tr>
<td>McSorley et al 2021</td>
<td>UK (Scotland)</td>
<td>SE</td>
<td>All</td>
<td>4841</td>
<td>HM-JACKarc</td>
<td>Colonoscopy</td>
<td>94.7% (91% to 97%)</td>
<td>4.7% (46% to 48%)</td>
<td>9.4% (8.4% to 10.6%)</td>
</tr>
<tr>
<td>Movat et al 2021</td>
<td>UK (Scotland)</td>
<td>SE</td>
<td>All</td>
<td>5381</td>
<td>HM-JACKarc</td>
<td>Various</td>
<td>86.7% (78.6% to 92.5%)</td>
<td>79.4% (78.3% to 80.5%)</td>
<td>7.7% (7.1% to 8.4%)</td>
</tr>
<tr>
<td>Pin Vieito et al 2021</td>
<td>Spain</td>
<td>SE</td>
<td>All</td>
<td>5623</td>
<td>OC-Sensor</td>
<td>Various</td>
<td>81.3% (71.3% to 88.3%)</td>
<td>84.1% (83.1% to 85.1%)</td>
<td>6.9% (5.4% to 8.7%)</td>
</tr>
<tr>
<td>Chapman et al 2021</td>
<td>UK (England)</td>
<td>DTA</td>
<td>NG12</td>
<td>732</td>
<td>OC-Sensor</td>
<td>Colonoscopy</td>
<td>89% (75% to 97%)</td>
<td>7.4% (70% to 77%)</td>
<td>16% (11% to 21%)</td>
</tr>
<tr>
<td>Farrugia et al 2021</td>
<td>UK (England)</td>
<td>DTA</td>
<td>NG12</td>
<td>519</td>
<td>HM-JACKarc</td>
<td>Colonoscopy / CTC</td>
<td>84% (69% to 94%)</td>
<td>78% (75% to 81%)</td>
<td>18% (12% to 24%)</td>
</tr>
<tr>
<td>D’Souza 2020</td>
<td>UK (England)</td>
<td>DTA</td>
<td>NG12</td>
<td>160</td>
<td>HM-JACKarc</td>
<td>Colonoscopy</td>
<td>87.5% (52.9% to 97.8%)</td>
<td>84.2% (77.6% to 89.2%)</td>
<td>22.6% (11.4% to 39.8%)</td>
</tr>
<tr>
<td>Khan et al 2020</td>
<td>UK (England)</td>
<td>DTA</td>
<td>NG12</td>
<td>928</td>
<td>HM-JACKarc</td>
<td>Various</td>
<td>85.1% (71% to 93.3%)</td>
<td>83.5% (80.8% to 85.8%)</td>
<td>22.6% (16% to 28.3%)</td>
</tr>
<tr>
<td>Nicholson et al 2020</td>
<td>UK (England)</td>
<td>SE</td>
<td>All</td>
<td>9896</td>
<td>HM-JACKarc</td>
<td>Record linkage</td>
<td>90.5% (84.9% to 96.1%)</td>
<td>91.3% (90.8% to 91.9%)</td>
<td>10.1% (8.15% to 12.0%)</td>
</tr>
<tr>
<td>Hernero et al 2020</td>
<td>Spain</td>
<td>SE</td>
<td>All</td>
<td>1572</td>
<td>OC-Sensor</td>
<td>Colonoscopy</td>
<td>93.5% (89.1% to 96.3%)</td>
<td>63.4% (60.7% to 66.0%)</td>
<td>28.9% (25.6% to 32.4%)</td>
</tr>
<tr>
<td>Widlak et al 2020</td>
<td>UK (England)</td>
<td>DTA</td>
<td>All</td>
<td>562</td>
<td>HM-JACKarc</td>
<td>Colonoscopy / CTC</td>
<td>80.0% (66% to 93%)</td>
<td>93% (91% to 95%)</td>
<td>44.0% (32% to 56%)</td>
</tr>
<tr>
<td>Arauz et al 2020</td>
<td>Spain</td>
<td>DTA</td>
<td>IDA</td>
<td>245</td>
<td>OC-Sensor</td>
<td>Colonoscopy</td>
<td>92.9% (76.5% to 99.1%)</td>
<td>57.1% (50.3% to 63.8%)</td>
<td>21.8% (15.4% to 30.1%)</td>
</tr>
<tr>
<td>Khasawneh et al 2020</td>
<td>UK (England)</td>
<td>DTA</td>
<td>CIBH</td>
<td>5818</td>
<td>OC-Sensor</td>
<td>CTC</td>
<td>88.6% (79.6% to 94.3%)</td>
<td>80.8% (79.7% to 81.8%)</td>
<td>5.3% (4.3% to 6.9%)</td>
</tr>
</tbody>
</table>

95% CIs given in brackets when reported.

CIBH, change in bowel habit; CRC, colorectal cancer; DTA, diagnostic test accuracy study; FIT, faecal immunochemical testing; IDA, iron deficiency anaemia; NPV, negative predictive value; PPV, positive predictive value; SE, service evaluation; SR, systematic review and meta-analysis.
Guidelines

92.0% (95% CI 84.4% to 95.9%) and specificity of 63.2% (95% CI 59.1% to 67.4%).85 The prevalence of CRC in this cohort with IDA, as expected increased with bands of fHb levels; 1.2% in those with fHb under 9 µg/g, 13.5% in the band 10–200 µg/g and 38.9% in those with cut-offs >200 µg/g. Another study in abstract form, suggested that fHb offered similar discriminatory values to symptoms and even younger patients.85

Previous BSG guidelines do not recommend use of FIT in those with IDA.86 Reasons for their recommendation (evidence strength low; statement strength weak) was related to publication bias in that iron deficiency anaemia may be over-represented in those with fHb below a threshold of 10 µg/g CRGs. However, the larger and subsequently published NICE FIT and qFIT studies more recently have provided further consistent data which supports the use of FIT testing in people with IDA.28 62

The meta-analysis informing the guidelines identified two studies (n=724) that reported specifically on IDA and used colonoscopy or CTC as a reference standard.61 The summary sensitivity at a cut-off of 10 µg/g was 96.7% (95% CI 88.7% to 99.6%) and specificity 73.6% (95% CI 70.1% to 76.9%).61

Change in bowel habit

Few studies have specifically reported on accuracy of FIT for CRC in patients with change in bowel habit (CIBH). Further, the definitions used make it more difficult to collate these together to form a unified consensus. In the Nottingham primary care study CIBH had lower CRC detection rates 4.5% versus 7.4% compared with those with IDA at a cut-off of 10 µg/g.74

The NICE FIT study suggested that the sensitivity of FIT for CRC at a cut-off of 10 µg/g in patients with CIBH, was higher in the older population (over 60 years of age) compared with those under 60 years of age (85.9% vs 60%), respectively.72

The large 38765 participants in the Spanish primary care study,87 reported broadly similar findings even when applying cut-offs of 10 or 20 µg/g/faeces.78 In fact, the number needed to scope for those presenting with diarrhoea was 15 and constipation 16.2 at 10 µg/g compared with 12.8 with diarrhoea and 13 with constipation at 20 µg/g.

The meta-analysis informing the guidelines identified two studies (n=10 067) reporting specifically on CIBH symptom and used colonoscopy or CTC as a reference standard. The summary sensitivity and specificity at 10 µg/g were 85.6% (95% CI 79.0% to 90.8%) and 83.6% (95% CI 82.9% to 84.3%), respectively.61

Evidence summary

1. FIT is highly sensitive for CRC in symptomatic patients with most large studies reporting a sensitivity of >87% at the commonly used cut-off of 10 µg/g.

2. The diagnostic accuracy of FIT is similar in both high and low risk symptomatic patients, irrespective of the cut-off used.

3. FIT is not always detectable in patients with rectal bleeding and is a useful evaluation tool when CRC is suspected. FIT is highly sensitive for CRC in patients with rectal bleeding with a sensitivity of >90% at a cut-off of 10 µg/g. Patients with negative FIT and persistent rectal bleeding can be safely investigated with flexible sigmoidoscopy and appropriate safety measures in place.

4. The evidence for use of FIT for the detection of CRC in IDA supports its use at a cut-off of 10 µg/g. If a lower cut-off is applied, then the accuracy further improves.

5. The evidence for use of FIT for the detection of CRC in those with isolated change in bowel habit is less clear although UK data suggests greater benefit in those over 60 years (while the Spanish data are consistent with other symptoms).

DIAGNOSTIC ACCURACY AND PATIENT-RELATED FACTORS

There is currently insufficient evidence to recommend variations in the fHb threshold for referral from primary care according to patient-related factors.

GRADE of evidence: low; Strength of recommendation: Strong.

In summary, age and gender may affect FIT performance but findings are inconsistent. There is no evidence to suggest that these are significant enough to warrant variations in thresholds at present—very young patients without genetic predisposition in whom the risk of CRC is very low may be the exception, however there is no evidence to exclude such individuals from FIT testing pathways currently. Therefore, in the absence of specific evidence to the contrary, the same fHb threshold should be used irrespective of patient-related factors (including: age-group, gender, ethnicity, deprivation and concurrent medication).

FIT, demographics and colorectal cancer

Demographic variations in CRC incidence are well recognised58 as are variations in fHb from screening studies.88 89 CRC diagnoses rise with increasing age and it is known that fHb also rises with age, even in the absence of notable pathology. CRC incidence is higher in men overall but also in the male population referred by GPs for further investigation. Deprivation has been noted to be a risk factor for men rather than women.88 A number of studies on FIT in symptomatic patients report higher fHb in men than in women, although Bailey et al’s study51 in a restricted DG30 population notes high fHb levels in women aged 30–40 years specifically. Men with CRC show a predilection to the rectum and one group has suggested false negatives in palpable rectal mass when a bleeding tumour would present overtly; CRCs in women are associated with the right colon and most studies report lower fHb in this group of cancers, most likely due to distribution of blood throughout a formed stool reducing concentration and increasing risk of sampling error. Although many studies show some differences in fHb by age and gender, findings are inconsistent, and this is perhaps unsurprising given the complexity of the potential interactions described and the size of cohort needed to address all of these. Furthermore, FIT is a test for occult blood in stool which may be related to many other pathologies with different interactions with demographics, or no identifiable pathology at all, adding further complexity to this challenge.

Age

A number of papers have reported data relevant to diagnostic performance of FIT by age. Interpretation is complicated by the different age categorisations, FIT thresholds and outcomes considered, as well as variations in the study populations and endpoints. Moreover, not all of the studies reported on diagnostic performance in terms of specificity, sensitivity and/or area under the curve (AUC).

A subanalysis of the NICE FIT study28 assessed diagnostic accuracy in 1103 patients under 50 years of age preselected by GPs for urgent referral to secondary care.90 All patients were well enough to undergo colonoscopy, thus potentially excluding some older and frailer patients. At all FIT thresholds, sensitivity for the older group exceeded that for the younger age group. At a threshold of 10 µg Hb/g faeces sensitivity was 87.5% for those <50 years of age versus 97.4% in those ≥50 years of age.
At the 2µg Hb/g faeces threshold, specificity was higher for the younger group (70.4% vs 64.1%); at the 10µg Hb/g faeces threshold, specificity was almost identical in the two groups (83.6% vs 83.5%); and at the 150µg Hb/g faeces threshold, it was slightly higher in the older group (92.2% vs 94.9%). At all thresholds, the PPV for the older groups was more than double that for the younger group (PPV was 6.8% vs 17.1% in those over 50y at 10µg Hb/g faeces). Despite these differences, further analysis of the 329 CRCs which were detected in this selected population found there was no association between age and FIT status (negative/positive) at a threshold of 2 or 10µg Hb/g faeces. The CRC prevalence in the younger age group was 1.5% and the study was not powered to assess younger patients specifically. Furthermore, the authors point out the value of detecting inflammatory bowel disease (IBD) and other pathology in younger patients.

The Fast Track FIT study of 5040 patients preselected by GPs for urgent referral, all completing either colonoscopy, CT colonography or CT abdomen and pelvis, compared groups above and below 60 years of age. AUC was slightly lower for those aged 60+ years of age (0.88, 0.85–0.92) compared with 1217 younger patients (0.92, 0.88–0.96), but the 95% CIs for these estimates overlapped. In this cohort, sensitivity and specificity were slightly higher in the younger group (sensitivity: 90.0 vs 83.5; specificity: 87.4 vs 85.4); PPV and NPV did not differ.

The authors describe an optimal cut-off threshold of 37 µg Hb/g faeces for those under 60 years of age, compared with a baseline of 20 µg Hb/g faeces for those over 60 years of age. In an earlier study this group also reported on 515 individuals similarly selected for investigation, in whom the FIT threshold to achieve optimal AUC for CRC detection again needed to be higher for younger patients (<65 years of age: FIT threshold ≥46 µg Hb/g faeces for AUC of 0.89 (0.722–1.000); 65+ years of age threshold ≥12 µg Hb/g faeces for AUC of 0.91 (0.842–0.981). However, these findings were based on only 7 cancers in the younger age group and 19 in the older group.

A small study of 404 patients, referred to secondary care for colonoscopy, in whom overall sensitivity and specificity of FIT at a threshold of 20 µg Hb/g faeces for CRC were 87.5% and 83.7%, respectively. FIT performance for relevant colonic pathology showed sensitivity of 50.6% and specificity of 69.6%; sensitivity was slightly lower and specificity slightly higher in those aged under 50, compared with those aged 50 and older. This study used a different FIT platform to most published studies (FOB Gold).

The largest study commenting on age reported on FIT in primary care use in Northern Spain and included 38 675 patients (CRC prevalence 1.7%) with a variety of symptoms. This study used 2 years registry-based follow-up as an endpoint and therefore the majority are not investigated in secondary care. However, the group reported diagnostic accuracy data based on follow-up and found no significant differences in sensitivity across three age strata:<50 years of age (93.1% at 10 µg Hb/g faeces; 91.8% at 20 µg Hb/g faeces), 50–69 years of age (91.5% at 10 µg Hb/g faeces; 88.7% at 20 µg Hb/g faeces) and >69 years of age (89.8% at 10 µg Hb/g faeces; 87.2% at 20 µg Hb/g faeces) although the values declined with increasing age at both cut-offs. They reported differences in specificity which fell significantly with rising age at a threshold of 10 µg Hb/g faeces (88.5% <50 years of age, 83.6% 50–69 years of age and 75% >69 years of age). In this study, the prevalence in the population under 50 years of age was only 0.3% and raising the threshold to 20 µg Hb/g faeces had the least detrimental effect on missed cancers in this age group; although the overall increase in missed CRC across the whole population at 20 µg Hb/g faeces was <1 in 1000.

Gender

Two large studies of FIT in high-risk cohorts selected for and completing secondary care investigation have reported diagnostic accuracy by gender. The NICE FIT study found no association between sex and FIT status using a cut-off of either ≥2 or ≥10 µg Hb/g faeces to define positivity. By contrast, the Fast Track FIT study found that, to achieve the optimum AUC for CRC, different FIT cut-offs would be needed for men and women; 21 µg Hb/g faeces for men and 16 µg Hb/g faeces for women gave AUC of 0.89 and 0.88, respectively. In their earlier, smaller, study of 515 patients, Turvill et al reported that the optimal FIT cut-off for detecting CRC was ≥22 µg Hb/g faeces for men (AUC = 0.909, 0.835–0.983) and ≥12 µg Hb/g faeces for women (AUC = 0.891, 0.744–1.000); but these results were based on 18 cases of CRC in men and 8 cases in women. In a study of 928 patients (41% men), FIT at a threshold of 10 µg Hb/g faeces had lower sensitivity for the detection of bowel disease (CRC, high-risk polyps or colitis) in women than in men. Sensitivity was 95.4% (95% CI 75.1% to 99.7%) for men compared with 76.1% (95% CI 54.5% to 89.9%) for women. Specificity was slightly higher in women (men: 80.5% (95% CI 75.9% to 84.4%); women: 85.5% (95% CI 82.2% to 88.4%).

In their large primary care study (n = 38 675) with follow-up, rather than full investigation, Pin Viento et al also reported higher sensitivity (91.6% vs 88.4%) but significantly lower specificity (79.9% vs 82.6%) in men at 10 µg Hb/g faeces. In a similar UK study of 9896 patients (41% men), where not all patients were investigated, Nicholson et al reported that the AUC for both CRC was almost identical in men (0.933) and women (0.948).

Age and gender

Nicholson et al also reported the AUC for FIT at a threshold of ≥10 µg Hb/g faeces for the detection of CRC if testing was restricted to different age groups. Restricting testing to people aged 80+ was the only instance where the AUC dropped below 0.90. If FIT testing, was to be restricted to those aged 40+, the AUC would be 0.944 (95% CI 0.899 to 0.988) for women and 0.934 (95% CI 0.897 to 0.972) for men; if limited to the 60+ age group, the AUC would be 0.919 (95% CI 0.856 to 0.982) for women and 0.921 (95% CI 0.870 to 0.971) for men; and if limited to those aged 70 and older, it would be 0.934 (95% CI 0.867 to 1.000) for women and 0.936 (95% CI 0.895 to 0.978) for men. In women, the AUC would be highest (0.708) when testing was restricted to those aged 50 and older; in men it would be highest if testing was restricted to those aged 70+.

Multivariate analyses including age and gender

A number of groups have reported multivariate analyses including demographics (with symptoms, blood results and other factors) as covariates and FIT measured by FIT is consistently the most predictive factor by some margin. No other factor reaches significance consistently, although increasing age and male gender appear most frequently.

The FAST score, combining age and gender with FIT, was developed in Northern Spain in patients undergoing colonoscopy, but has failed to show improved clinical effectiveness compared with FIT alone when trialled in a broader population in Tayside and when applied to the NICE FIT data set. It is not clear whether the use of different platforms affected these findings, as well as other intrinsic factors.
differences in the cohorts evaluated. The ColonFlag score takes this approach a step further, combining FBC results with Hb, age and sex, and showed improved specificity compared with FIT—taking an AND/OR approach sensitivity improved to 100%. However, this study is relatively small and needs further validation. A similar sized study (n=408) by Digby et al found no value in other factors, including demographics, other than family history.

A much larger study by Withrow et al reported on 16604 patients tested in primary care (and includes the cohort from Nicholson et al) specifically focussing on the value of combining FIT with blood tests. However, the majority of patients were not investigated. In a variety of models, they found no value in including age. They also reject the value of models including gender, although they note that the threshold to reach a 3% PPV is slightly higher in men (25 µg Hb/g faeces) than in women (17 µg Hb/g faeces). Consistent with this, Rodriguez-Alonso et al found that, after accounting for FIT, sex was statistically significantly associated with both CRC and advanced neoplasia and men had more than twofold increased risk of disease. They found age did not need to be included in a multivariable model for CRC, after FIT (and gender and IDA) had been included. However, when an outcome of advanced neoplasia was considered in a model including FIT result and gender, risk increased with increasing age.

There is insufficient evidence that diagnostic performance of FIT for the detection of colorectal neoplasia varies by ethnicity, deprivation and other factors.

Ethnicity and deprivation

Although some studies provide breakdown of populations by ethnicity and deprivation, none have looked at this as a primary outcome and there is no clear data on variations in diagnostic accuracy. Even in larger data sets the low event rate in non-white categories or specific deprivation groups appears too small and there are no pooled analyses that address this. The NICE FIT study compared the characteristics of 329 patients detected with CRC, all of who had undergone a FIT, were compared according to notional FIT status. At a positivity threshold of either ≥2 µg or ≥10 Hb/g of faeces, there was no statistically significant difference in the ethnic distribution of those who would have been classified FIT positive or FIT negative. The same observation was made for deprivation category of area of residence. A questionnaire based evaluation of the NICE FIT cohort suggests lower acceptability in non-white ethnic groups.

The group in Nottingham have also reported some evidence that the non-return of FIT kits is associated with younger age, male sex, non-white ethnicity and higher deprivation populations (Abstract added to library). These findings mirror studies in screening and have implications for education, implementation and safety netting, rather than diagnostic accuracy per se.

Other factors

As noted earlier, Digby et al reported on family history and after adjusting for FIT result (and rectal bleeding and folate level) a family history of polyps was associated with a more than eightfold increased risk of any significant bowel disease (CRC, advanced adenoma (AA) or IBD) (OR=8.21, 95%CI 1.74 to 38.78). Only nine people with a family history of polyps had significant bowel disease.

A single study has examined associations between smoking and body mass index and risk of advanced colorectal neoplasia, once FIT had been taken into account. There was a significant association between being a current or ex-smoker and risk of advanced colorectal neoplasia (ACN) (multivariate OR=1.51, 95% CI 1.02 to 2.29). Body mass index greater than 25 kg/m² was not significantly associated with ACN.

Evidence is too limited to conclude whether the diagnostic performance of FIT varies in people using specific medications.

Medications

In total five publications have reported on diagnostic accuracy of FIT in patients using particular medications. Two studies examined proton pump inhibitor (PPI) use. Rodriguez-Alonso et al included 1002 patients referred for colonoscopy; 40% were PPI users. In total 133 patients had advanced neoplasia (AN) and 30 patients had CRC. There were no differences in sensitivity or specificity of FIT with a positivity threshold of ≥20 µg Hb/g of faeces for the detection of CRC among those who used PPIs and those who did not. When AN was considered, both sensitivity and specificity were significantly lower for PPI users (sensitivity: 43.0%; specificity: 86.9%) than non-users (sensitivity: 65.6%, p=0.009; specificity: 92.3%, p=0.010). The second, smaller, study of 612 patients published only as an abstract, reported that sensitivity of FIT at a threshold of 10 µg/g for advanced neoplasia (n=55) was lower in PPI users than non-users (54% vs 81%, p=0.05), while specificity did not differ. The area under the receiver operating characteristic curve was 74% (93% CI 0.58 to 0.91) for PPI users compared with 0.92 (95% CI 0.89 to 0.95) for non-users.

Three publications considered the possible influence of use of antiplatelet or anticoagulant medication on FIT diagnostic accuracy. Two reported on data from the COLONPREDICT study; the first, an abstract from 2014, included 1567 patients; the second, a full paper published in 2018, included 3052 patients. The smaller study reported that the diagnostic accuracy of FIT for CRC was significantly lower among those taking antiplatelet and/or anticoagulant medication (AUC: users 0.81; non-users, 0.88; p=0.04). The larger study focused on aspirin specifically in a study in which 16% of patients used aspirin. Continuous treatment with aspirin did not influence sensitivity, specificity of the AUC of FIT for either CRC detection or AN detection at a threshold for positivity of ≥20 µg of Hb/g of faeces. In a subgroup analysis of patients using ≥300 mg/day aspirin, sensitivity, specificity and AUC were lower, as was polyp prevalence, than among aspirin non-users, but this group included only 58 people and the differences were not statistically significant. The final study reported only multivariable modelling results and use of anti-coagulants, anti-platelets or non-steroidal anti-inflammatory drugs (NSAIDs) was not significantly associated with advanced colorectal neoplasia after adjusting for FIT results.

There is currently insufficient evidence to confirm whether diagnostic accuracy is impacted by the type of FIT analyser used.

GRADE of evidence: low; Strength of recommendation: weak

There is no international standardisation of FIT methods meaning that different results could be obtained on different manufacturer systems. Despite this, in symptomatic testing, single thresholds for referral have been recommended.

To date there are only two peer-reviewed publications that have directly compared results obtained on two different FIT analytical systems when patients have taken samples from the same bowel motion in a symptomatic pathway.

In the first study, 732 patients returned both an OC-Sensor and an HM-JACKarc collection device. They had been instructed to collect samples into each device from the same bowel motion. Correlation of results was carried out and agreement at cut-offs of 4, 10 and 150 µg Hb/g faeces (µg/g) were assessed. To act as a...
control 114 patients collected two samples using two OC-Sensor devices. At thresholds of 4, 10 and 150 µg/g the Cohen’s kappa have concluded that there is not enough data to comment on the comparative performance is 0.74, 0.79 and 0.76, respectively, which is interpreted as substantial agreement. When two OC-Sensor devices are compared the Cohen’s kappa are 0.80 (substantial agreement), 0.91 (almost perfect agreement) and 1.00 (almost perfect agreement) respectively. This suggests that the referral rate will vary dependent on which of the two methods is used.

In terms of diagnostic accuracy, at thresholds of 4, 10 and 150 µg/g, OC-sensor had a higher sensitivity and lower specificity than HM-JACKarc for CRC. Thus, based on this study, employing the same thresholds for the two methods OC-Sensor will generate more referrals however it will also detect more CRCs than HM-JACKarc at the same thresholds.

In the second study, the QuikRead Go (QRG), a quantitative point of care FIT test which had previously undergone independent analytical evaluation, was compared with the FOB Gold wide method on the SENTiFIT laboratory analyser. Five hundred and fifty-three patients provided paired samples for both methods and underwent colonic investigations that were suitable to give definitive diagnostic outcomes. Fourteen patients were diagnosed with CRC. QRG reported one false negative. FOB Gold reported no false negatives. Thirty per cent of QRG results were >10 µg/g would have resulted in referral compared with 16.9% for FOB Gold wide.

In an unpublished study, 233 patient returned FIT devices from four different FIT systems collected from a single bowel motion. To act as a control a further 189 patients returned two FIT devices from the same FIT system. Differences were observed in the referral rates for different methods and the categorisation according to Cohen’s kappa, specifically for one method more than the other three. There were only seven CRCs detected in the four FIT group so inadequate data to comment conclusively on the comparative diagnostic accuracy of the different FIT tests.

In addition to the three studies above, four systematic reviews have commented on the different FIT assays available. The conclusions were that there are a lack of studies directly comparing the performance of different FIT assays and that there are currently no data on the comparative performance of different FIT assays. In addition, it was reported that the limited number of studies, the majority of which were using OC-Sensor, along with high study heterogeneity, did not enable conclusions to be drawn from combining data from different studies.

FIT COMBINED WITH OTHER FACTORS TO OPTIMISE RISK STRATIFICATION

There is currently insufficient evidence to recommend including FIT in a risk score with other clinical features to identify patients with symptoms of suspected CRC.

GRADE of evidence: low; Strength of recommendation: weak.

There is some supporting and emerging evidence that combining fHb with either a composite score or another biomarker, improves CRC detection. However, these methods have not yet been clinically validated.

Several scoring systems have been reported either in combination with fHb or in combination with fHb alone to improve detection of CRC. The FAST score devised initially by the Spanish group uses a combination of age, sex, fHb at different cut-offs. As expected, when using a lower FAST score cut-off (>2.12) provided almost 100% sensitivity with poor specificity 14% and a 100% NPV. For detection of advanced neoplasia in a British population, there was a 10% improvement compared with fHb on its own. By contrast, Digby et al did not show any benefit of using the FAST score compared with fHb on its own. The latter was in a primary care setting where a quarter had a colonoscopy for final diagnosis.

FAST score also performed less well against ColonFlag (an Israeli trademarked algorithm) that comprises FBC, red cell indices, ferritin, iron and transferrin. CRC accuracy was reported to provide sensitivity of 100% compared with 73% with FAST score. Specificity was poor at 50% with ColonFlag but 81% with FAST score and NPVs of 99% and 100% for FAST score and ColonFlag, respectively. Using a lower of two proposed cut-offs for ColonFlag, CRC accuracy resulted in a sensitivity of 80%, specificity of 48% and NPV of 99% (sample size was limited to 21 cases).

ColonPredict which used a combination of symptoms, fHb, serum haemoglobin and mean cell volume was deemed superior to symptoms alone while ColonOFIT, which uses three serial fHb measurements in a week, patient questionnaire, mediation consumption (eg, NSAIDs) had a ninefold higher OR of detecting CRC than serial fHb on its own. The RAT (research assessment tool) which comprises clinical, demographic, fHb, blood markers and colonoscopy outcome seems to hold promise with diagnosis of bowel disease (defined by authors as CRC and significant adenoma but excluding IBD). The OR was 9 (4.3–18.6) using the RAT tool compared with 5.3 (2.4–11.7) with fHb on its own (Lord et al, 2018). The Health Technology Assessment from 2017 using 10 studies (including 9 from secondary care), demonstrated that fHb on its own was still more effective and cost-effective compared with faecal occult blood testing or using no triage test.

Volatile organic compounds (biomarkers of cellular inflammation and/or cancer) have shown some promise with Widlak et al showing improved CRC detection in those who are tested negative with fHb (<10 µg/g faeces) in terms of improving its sensitivity from 80% to 97%. A recent network meta-analysis of both fHb and volatile organic compounds for CRC detection demonstrated the probability of CRC detection improving from 0.5% to 0.1% when both tests were negative.

FIT IN SPECIFIC POPULATIONS

We suggest that FIT may be used to stratify adult patients aged younger than 50 years with bowel symptoms suspicious of a diagnosis of CRC.

GRADE of evidence: low; Strength of recommendation: weak.

The incidence of CRC in younger patients under the age of 50, also known as early onset CRC, has been documented in studies across developed healthcare economies. Recent studies suggest incidence in this age group is rising. We suggest that FIT may be used to stratify adult patients aged younger than 50 years with bowel symptoms suspicious of a diagnosis of CRC referred from primary care for further investigation, at the same threshold as for older patients. CRC incidence will be low in younger patients at ‘low’ FIT thresholds, but FIT would be of value to detect other serious bowel disease in this group.

CRC can be difficult to detect on the basis of symptoms in younger patients, as these may frequently overlap with common benign conditions. For example, change in bowel habit or abdominal pain may be due to irritable bowel syndrome, and rectal bleeding is frequently caused by haemorrhoids in younger patients. A diagnostic test that can identify those younger patients at risk of cancer may therefore be useful as an adjunct to decide on referral to secondary care for further investigation.
The diagnostic accuracy data of FIT in younger patients (<50) was the primary endpoint in one study, but reported in subgroup analyses in three further studies. Souza et al investigated the diagnostic accuracy of FIT in 9822 symptomatic patients in the UK, and had a particular focus on 1103 symptomatic patients under the age of 50. The prevalence of CRC was 1.5% (16/1103) in younger symptomatic patients. The sensitivity of FIT for younger patients aged <50 was 87.5% (95% CI 61.7% to 98.4%), 81.3% (95% CI 54.4% to 96.0%) and 68.8% (95% CI 41.3% to 89.0%) at fHb cut-offs of 2 and 150 µg/g, respectively. Specificity at these cut-offs was 70.4% (95% CI 67.6% to 73.1%), 83.6% (95% CI 81.3% to 85.5%) and 92.2% (95% CI 90.4% to 93.7%). At each threshold, the sensitivity was lower for younger patients than for patients aged 50 and older, but specificity was higher, and the differential between age-groups narrowed as the threshold rose. In those under 50 years of age, the PPV for CRC increased from 4.2% (95% CI 2.3% to 6.9%) to 11.5% (95% CI 5.9% to 19.6%) at cut-offs of 2 and 150 µg/g. The higher prevalence of IBD in younger patients meant that the PPV of FIT for serious bowel disease (CRC, IBD and AA) was high, increasing from 31.3% (95% CI 26.3% to 36.5%) to 63.6% (95% CI 55.2% to 75.0%) at the same cut-offs.

In a study in Denmark, of FIT in patients over the age of 30 with low-risk symptoms the prevalence of CRC was low at 0.5% (4/848) in the subgroup of patients under the age of 50 and no cancers were detected in those aged under 40. The PPV of FIT (>10 µg/g) for CRC in the 40–49 age-group was only 0.6% (0.1%–1.3%). In a study from Spain, Lue et al reported diagnostic performance of FIT at a threshold of 20 µg/g for the detection of any relevant colonic pathology (CRC, AA, IBD, microscopic colitis or angiodysplasia). In the subgroup aged under 50 (n=119), specificity and NPV exceeded 90% (specificity: 92%; NPV: 90.2%) but sensitivity was only 47.4% and PPV was 52.9%. Finally, in a mixed population of 38,675 asymptomatic and symptomatic patients (8866 aged <50), not all of whom underwent colonoscopy or other diagnostic investigation, Pin Vieito et al found that sensitivity and specificity of FIT at a threshold of 10 µg/g for the detection of CRC up to 2 years later in those aged under 50 were 93.1% (95% CI 78.0% to 98.1%) and 88.5% (95% CI 87.9% to 89.2%), respectively. PPV in this younger age-group was 2.6% (1.8%–3.8%). The authors reported that sensitivity did not vary by age, but that specificity was lower, and PPV was higher, in older patients (50–69 and 70+) than in those under 50. The finding that right-sided CRC are more likely to be missed by FIT suggests that younger patients, who are more likely to present with distal left-sided CRC, may not be at increased risk of false negative FIT.

In summary, FIT can be used to risk stratify the risk of CRC or serious bowel disease in younger patients (aged <50). The prevalence of CRC in symptomatic patients and the sensitivity and PPV of a positive FIT, is lower in younger than older patients. However, other serious bowel conditions may cause a positive FIT and merit investigation, particularly when higher fHb concentrations are detected. Further research is needed to confirm the diagnostic accuracy of FIT specifically in younger patients, the optimal FIT threshold and whether testing should be limited to those older than a specified age (eg, 40 years); the relative costs and benefits (in terms of detection of both CRC and other colonic pathology) of different strategies are not clearly established.

**INVESTIGATION IN SECONDARY CARE**

Colonoscopy is considered the standard method of investigation, however other methods of colorectal imaging may be appropriate in some patients.

**GRADE of evidence: low; Strength of recommendation: weak.**

We recommend that for patients with symptoms of a suspected diagnosis of CRC, CTC is equivalent to colonoscopy for detection of CRC (the choice of modality should be determined by the local expertise and availability). **GRADE of evidence: low; Strength of recommendation: Strong.**

There is currently insufficient evidence to support use of a specific quantitative FIT threshold to recommend the selection of CTC versus colonoscopy. **GRADE of evidence: very low; Strength of recommendation: weak.**

In this section we considered the evidence for colorectal investigation in patients with signs or symptoms of a suspected diagnosis of CRC, however there is limited available direct evidence in this specific population with a FIT above threshold with most studies reporting colonoscopic outcomes as a reference standard. In the UK NHS, only colonoscopy and CTC are established appropriate whole colon investigations for the exclusion of CRC and large polyps in patients with symptoms suggestive of CRC. CTC has superseded barium enema examinations. Colonoscopy in the UK is quality assured by the Joint Advisory Group for GI endoscopy, which has led to steady improvements in service quality, enables biopsies to be taken at index investigation and more effective diagnosis of non-neoplastic pathology.

Historically colonoscopy was the criterion standard for lower GI investigation allowing both direct visualisation and biopsy or polypectomy in a single procedure; however from 2005, NICE guidance supports the use of CTC as an alternative test to colonoscopy with adequate evidence on safety and efficacy. Guidance from European Society of Gastrointestinal Endoscopy/European Society of Gastrointestinal and Abdominal Radiology (ESGE/ESGAR) in 2020 recommend CTC as an acceptable and equally sensitive alternative for patients with symptoms suggestive of CRC when colonoscopy is contraindicated or not possible (strong recommendation, high quality evidence). Because of lack of direct evidence, ESGE/ESGAR did not recommend use of colon capsule endoscopy (CCE) in this situation (very low quality evidence). CTC is safe with complications rarely encountered and rarely serious. CTC is well tolerated even when colonoscopy is contraindicated or incomplete and frequently used in older patients.

The SIGGAR study, a landmark multicentre randomised study of 1610 patients with symptoms suggestive of CRC from 21 UK NHS hospitals showed:

1. Detection rates for large polyps or CRC colonoscopy and CTC were the same at 11%.
2. Referral rates for additional colonic investigation was 30% for CTC and 8% for colonoscopy.
3. 10% of patients had a significant extracolonic finding at CTC: 2% had extracolonic malignancy and 3.5% had an extracolonic diagnosis that at least part explained presenting symptoms.
4. Incompletion rate for CTC was 4% versus 7% for colonoscopy.

The SIGGAR study also showed that the patient acceptability and psychological impact of investigation via CTC or colonoscopy was similar.
A systematic review and meta-analysis by Obaro et al. of interval cancer rates (post imaging CRC rates) for CTC and colonoscopy have been published recently. A systematic review and meta-analysis by Obaro et al. showed a post CTC CRC rate of 4% (4.4 missed cancers per 100 detected, with low heterogeneity) which is within the range reported for colonoscopy (3%-9%). Systematic review and meta-analysis of CTC after positive faecal occult blood test or FIT showed CTC had high per-patient sensitivity (89%) for 6 mm+ lesions (adenomas or cancer) and pooled sensitivity for cancer of 96% (low heterogeneity). Specificity for polyps (6 mm+) was lower (75%) and heterogeneous with performance contingent on centre.

CTC has a potential advantage over other whole colon investigations by also detecting extracolonic malignancy or other life-threatening conditions such as symptomatic abdominal aortic aneurysm which may be responsible for symptoms. A 2WW pathway audit of 1792 straight to CTC patients identified non-colonic cancer in 4.3% patients and 12% had a new, potentially significant extracolonic finding. A 2018 systematic review and meta-analysis (44 studies included from screening and symptomatic populations) showed potentially significant extracolonic findings in 5.2% of symptomatic individuals. The rate increased to (5.7%) for those aged over 65 years versus 2.3% for those younger than 65 years with an 8% overall referral rate for additional investigation of these findings. Concerns about ‘over investigation’ of false positive or unimportant findings following CTC appear overstated as patients and their referring clinicians are prepared to accept a much higher rate of additional investigations (in up to 100% and 40% of examinations, respectively) than occurs in real-world practice.

There is no evidence for selecting one whole colon examination over another on the basis of FIT level in terms of diagnostic accuracy or enhanced patient pathway. While intuitively beneficial, there is a lack of evidence recommending use of CTC for patients with symptoms of abdominal pain and weight loss. Similarly, there is no specific evidence favouring CTC over colonoscopy or vice versa below a specific FIT threshold, despite a relative increase in the likelihood of extracolonic pathology with lower FIT levels. A 2007 study showed symptomatic patients with no colonic abnormality are more likely to have an important extracolonic finding and extracolonic findings could account for 10% of patients’ symptoms at initial presentation.

CTC does not offer immediate biopsy or polyp resection with even the most experienced centres referring approximately 10% of cases to endoscopy, potentially slowing patient pathways, although these levels will be considerably higher in a ‘FIT above threshold’ population.

The small dose of radiation administered with CT colonography can be minimised with newer state-of-the-art CT platforms and protocols (equivalent to approximately 1-3 years of background radiation in the UK), including routine use of dose modulation and ultra-low dose protocols for younger patients. However, as a general rule, radiation associated tests should be used judiciously according to assessment of risk/benefit as governed by IRMER (Ionising Radiation (Medical Exposure) Regulations) with particular caution in younger patients.

Currently there is an NHS CCE trial which will provide data on the utility of this intervention in the investigation of this patient population, however to date there is no direct evidence in the symptomatic population. Early indirect evidence from the SCOTCAP study suggest that CCE may be well-tolerated, although that there is a high rate of incomplete examination. There is a suggestion CCE may be considered in some lower risk populations whereby colonoscopy might be avoided. In due course the NHS England trial will provide further relevant data to inform the use of CCE in the near future, however currently there is no published data of the diagnostic accuracy of CCE in people with symptoms of a suspected diagnosis of CRC.

ACCEPTABILITY

On the basis of limited evidence, clinicians and patients consider FIT as an acceptable test for symptomatic CRC in most circumstances

GRADE of evidence: very low; Strength of recommendation: weak

We recommend that services should consider ways of promoting a high proportion of patients to return a FIT kits.

GRADE of evidence: very low; Strength of recommendation: Strong.

In summary, studies which report the uptake in symptomatic populations demonstrate an uptake of between 78.9% and 94%. There was some suggestion that younger age groups found the FIT kits less acceptable to complete which may need to be explored in more detail and addressed. These results suggest that the test has a high degree of acceptability. Patients prefer the non-invasive FIT kits over colonoscopy as long as its accuracy is comparable, however GPs’ acceptance prior to COVID-19 was more limited. More resources need to be invested when publishing these new guidelines with audit and evaluation demonstrating that all areas of the country are delivering an equitable service.

In this systematic review observational or qualitative studies were identified including those which reviewed FIT kit response rates and the practice of the referring clinicians. There is little ‘direct’ evidence of acceptability studies of the use of FIT in symptomatic populations.

Von et al. surveyed 1057 adults aged between 40 and 59 to imagine they had symptoms of CRC and answer a survey exploring their choice of FIT versus colonoscopy as a diagnostic test. Potential ‘miss rates’ of CRC were suggested. Interestingly 150 chose neither test, while 70% chose FIT when the miss rate was equivocal in both FIT and colonoscopy but when the miss rate was increased by one person in the FIT group the acceptance reduced to 40.4%. Information of a normal result by letter was preferred by 62.2% of the patients while 32% wanted face-to-face appointment to discuss abnormal result and 7.1% would still chose FIT even with a 10% reduction in accuracy compared with colonoscopy. One-third of GPs preferred to use FIT to ‘rule out’ colonoscopy but there was confusion over symptoms and optimal use of FIT. The majority of GPs were still not using FIT routinely at the time of the survey.

Digby et al. report that of 4072 FIT who were sent by secondary care to patients presenting with lower bowel symptoms, 2881 returned their FIT kits, suggesting a return rate of 70.75%.

In a service evaluation of FIT and anaemia for risk stratification in the 2-week pathway for CRC, 1106 FIT kits were sent with a return rate of 80.9%.

Maclean et al. asked 381 symptomatic patients to undergo FIT with 358 (94%) samples were returned. Onward referral for colonoscopy reduced from 62% to 34%. Follow-up of all the patients over a 2-month period found one person who had returned a positive FIT but had declined investigation because of fears over COVID-19, had later been diagnosed with CRC.
Ng et al stated that 17 out of 19 patients (89.5%) referred by the GP on ‘non-cancer pathways’ and 348 out of 441 referred into the urgent cancer pathway (78.9%) had returned their FIT kits within 7 days and overall 94.4% returned their kits with 14 days.137

Chapman reviewed 1862 patients referred to their GP with lower GI symptoms, 91.4% returned their kits within 7 days, however the authors noted that those not returning their FIT kits were significantly younger than those who did.

A recent qualitative survey of symptomatic patients assessed usability and acceptability of FIT and reviewed 1151 patients who had reported symptoms of CRC and sent FIT tests.40 A relatively high 90.2% found the kits straightforward to use, 76.3% disagreed that the tests were unhygienic and 78.1% preferred FIT to colonoscopy.

In 2018, Von et al explored the attitudes of GPs towards FIT in patients at an increased risk of CRC.138 One-third preferred to use FIT as a ‘rule out’ test. They were more willing to use FIT if the GPs were aged between 36 and 45, considered FIT to be highly accurate, thought the patient would benefit FIT over immediate colonoscopy, and were highly confident about discussing FIT tests. GPs were also less willing to offer FIT if they referred more than 10 patients onto the 2WW pathway per year or thought the patient needed a longer consultation for FIT. This paper suggests the acceptance of FIT by clinicians was still low at the time of publishing and that any changes to the testing for CRC in symptomatic patients may facilitate access for some groups who were previously unwilling to come forwards for investigation of their lower GI symptoms due to concerns about invasive testing.

For those with dexterity difficulties in performing the test, digital rectal examination to obtain stool for FIT testing appears to offer similar accuracy to home performed tests.79 This might be combined with point-of-care testing to allow for discussion of an appropriate whole colon examination following a positive test with the clinician to support engagement with downstream testing for CRC in a single consultation.143

**DISCRIMINATION**

We recommend that clinicians actively prevent discrimination at any stage of the diagnostic pathway as symptomatic FIT testing is rolled out, with a focus on equity of access and application to all patients with lower GI symptoms.

**GRADE of evidence: very low; Strength of recommendation: Strong.**

While the data on discrimination in symptomatic FIT are extremely limited, active efforts should be made to avoid discrimination as symptomatic FIT testing is rolled out, with a clear emphasis on equity of access and application.

Data on the role of underutilisation of CRC screening among certain racial and ethnic minorities, age groups and among persons with lower socioeconomic status in the screening literature are well reported139; however, data on differences in utilisation for FIT testing in symptomatic patients is very limited. Differential FIT utilisation can occur for a range of reasons: due to inability to perform the test, for example, due to rheumatological or neurological disability preventing fine motor skills to collect the sample; blindness; unwillingness to engage with stool based testing, perhaps due to level of disgust in performing the test, with some evidence from screening that more disgust sensitive individuals may be disinclined to complete any test involving collection of faces140; and subsequent unwillingness to proceed to whole colon examination after a positive FIT result. In an US screening observational study only 43% of FIT positive patient completed colonoscopy by 6 months.141 Nine per cent (103/1083) of patients with a positive FIT test for CRC symptoms did not proceed to secondary care assessment within 28 days of the test in a study in the NHS (Mr M Abulafi, personal communication 3rd May 2022).

It is also possible that clinicians may choose to use FIT differentially in patients presenting with lower GI symptoms based on their assessment of the pre-test probability for CRC for the patient in front of them, a conscious bias, or equally if GPs feel less confident about discussing the benefits of FIT with patients (OR 2.15, 95% CI 1.46 to 3.16)138; however there remains a risk of unconscious bias. If patients perceive medical discrimination, they may be less likely to come forward for screening. In a Californian cohort, women perceiving medical discrimination (racial or ethnic-based) were less likely to be screened for CRC (OR, 0.66; 95% CI 0.64 to 0.69),142 which might also reduce engagement with FIT-based symptomatic testing.

Translating the data from screening, where the whole population are invited for a test, to scenarios where patients are seeking healthcare for lower GI symptoms is a challenge, which may be exacerbated by whether the assessment is made in primary or secondary care. However interestingly from the screening literature, differences in choice of test occurred over time following changes in CRC screening options by the US Preventive Services Task Force, where when less invasive screening options were available (FIT and multitarget stool DNA tests), they saw increased use of less invasive options.141 The new widespread availability and promotion of FIT-based testing for symptomatic patients may facilitate access for some groups who were previously unwilling to come forwards for investigation of their lower GI symptoms due to concerns about invasive testing.

For those with dexterity difficulties in performing the test, digital rectal examination to obtain stool for FIT testing appears to offer similar accuracy to home performed tests.79 This might be combined with point-of-care testing to allow for discussion of an appropriate whole colon examination following a positive test with the clinician to support engagement with downstream testing for CRC in a single consultation.143

**IMPLEMENTATION**

We recommend that FIT, as a diagnostic triage tool, can be implemented safely at primary care level, and that a programme of education be developed to facilitate implementation of FIT in primary care.

**GRADE of evidence: very low; Strength of recommendation: Strong.**

Most studies of the use FIT in the symptomatic population are essentially studies of diagnostic accuracy. However, FIT is not a diagnostic test for colorectal neoplasia, and even at very low thresholds, sensitivity is not 100%, that is, not all cancers will be detected. Thus, for it to act as an effective means of triaging symptomatic patients for further diagnostic investigation, it is essential that it should be employed as an aid to decision-making against a background of clinical acumen and auxiliary tests, especially an FBC. Furthermore, in order that the impact on referral to secondary care and subsequent diagnostic workload is maximised, it can be argued that the ideal stage in the patient pathway to use FIT is in primary care.

Currently there is a paucity of pragmatic implementation studies of FIT as a diagnostic aid with a view to optimising referral patterns. However, there are some studies that provide useful data.

Mowat and his colleagues16 have reported on the outcomes of a service development where GPs were encouraged to use FIT in addition to clinical assessment and FBC irrespective of symptoms. The fHb was measured using HM-JACKarc (Kyowa
Medex) with a recommended cut-off of ≥10 \( \mu g \) Hb/g faeces. Anonymised record linkage to the Scottish Cancer Registry was used to find incident cases of CRC. During the study period FIT specimen were submitted for 5422 patients and the positivity at the chosen threshold was 21.9%. Irrespective of FIT result, 2848 patients had an immediate referral to secondary care and 3 with Hb < 10 \( \mu g/g \) presented with obstructing CRC shortly after submission of the FIT. Colonoscopy was carried out in 1447 and the prevalence of SBD was 20.5% (95 CRC (6.6%), 133 high-risk adenomas (9.2%) and 68 IBD (4.7%)); this represented 6.6% of patient with an Hb < 10 \( \mu g/g \) and 32.3% in those with fHb ≥ 10 \( \mu g/g \). There was no immediate referral in 2521 patients 95.3% of whom had Hb < 10 \( \mu g/g \). Four of these (0.2%) were later diagnosed with CRC. The record linkage did not identify any additional CRC cases within a follow-up period of 23–35 months. In the first year of this service, a reduction in referrals of 15.1% was seen.

McSorley and others \(^1\) reported data from three Scottish NHS Boards, where HM-JACKarc FIT kits were employed by GPs as an aid to referral in the same way as Mowat et al. \(^1\) In total 4840 patients who had colonoscopy after FIT submission were included. Of 2166 patients (44.7%) with Hb < 10 \( \mu g/g \) faeces (\( \mu g/g \)), 14 (0.6%) had a diagnosis of CRC, with the NNS of 155. In the 2675 patients (55.3%) with Hb ≥10 \( \mu g/g \), there were 252 CRCs found (9.4%) with an NNS of 11. In 705 patients with fHb ≥400 \( \mu g/g \), 158 (22.4%) had CRC with an NNS of 5. More than 50% of those with CRC and an fHb <10 \( \mu g/g \) had coexisting anaemia.

Chapman et al. \(^1\) incorporated postal FIT into the CRC 2WW pathway in those without rectal bleeding for a 1-year period. A total of 1106 patients received FIT and 80.9% returned them; 810 patients were investigated and 40 CRCs were found (4.9%). 60.4% of all patients had a FIT result lower than 4 \( \mu g/g \), and 69.7% had a result of <10. Sixty per cent of patients with CRC had a FIT result of ≥150 \( \mu g/g \). In five CRCs in patients with a FIT value < 10 \( \mu g/g \) there was either anaemia or a palpable rectal mass or the patient was anaemic. A FIT result > 10 \( \mu g/g \) was associated with a 97.5% sensitivity and 64.5% specificity for CRC whereas a result >4 \( \mu g/g \) and/or anaemia was 100% sensitive and 45.3% specific for CRC.

Bailey et al. \(^1\) carried out a service evaluation of GP access to FIT as an aid to CRC diagnosis. Over a 1-year period, there were 5733 FIT results, of which 4082 (71.2%) were <4.0 \( \mu g/g \), 579 (10.1%) were 4.0–9.9 \( \mu g/g \), 836 (14.6%) were 10.0–149.9 \( \mu g/g \) and 236 (4.1%) were >150.0 \( \mu g/g \). A 33% rise in urgent referrals was seen during the evaluation. In the 4082 patients with FIT <4.0 \( \mu g/g \) two CRCs were diagnoses. Of all the CRCs 58.4% of those associated with a positive FIT result were early (Stage I and II) and the percentage of CRC diagnoses that arose from an urgent referral rose after introduction of FIT.

Finally, in a study of GPs’ attitudes to FIT, Von and his colleagues, \(^1\) conducted prior to higher FIT utilisation during the COVID-19 pandemic, found that only one-third of GPs would prefer to use FIT rather than the current 2WW criteria for referral.

Anecdotal experience from the Thames Valley region, where FIT was rolled out rapidly due to the COVID-19 pandemic suggests that effective communication between clinical commissioning group and cancer alliances with GPs is critical, via webinars and more traditional methods, for example, newsheets. Published evidence was important for many GPs to make practice change, as was the availability formal advice from NHSE (National Health Service England). Oxford was an early adopter of FIT in 2016 and may not be representative as many GPs were already using this technique. When approaching regions that did not have locally available FIT testing embedded, there was considerably more resistance to the move. Clarity on the mechanisms and role of safety netting was also critical in driving engagement. In addition, standardisation of reporting which includes clear instructions about the clinical actionability of FIT results would facilitate interpretation in primary or secondary care.

Further advice about implementation with signposting to existing programmes in available in online supplemental file 2.

### Author affiliations

1. The Wolfson Endoscopy Unit, Gastroenterology Department, St Mark’s Hospital and Academic Institute, Harrow, UK
2. Faculty of Medicine, Department of Surgery & Cancer, Imperial College, London, UK
3. Department of Colorectal Surgery, University Hospital of Wales, Cardiff, UK
4. Colorectal Surgery, Croydon Health Services NHS Trust, Croydon, UK
5. Nottingham Colorectal Service, Nottingham University Hospitals NHS Trust, Nottingham, UK
6. Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
7. University of Warwick, Clinical Sciences Research Institute, Coventry, UK
8. Gastroenterology Department, University Hospital Coventry, Coventry, UK
9. Patient Representative, London, UK
10. Hub Director, NHS bowel Cancer Screening South of England Hub, Royal Surrey County Hospital NHS Foundation Trust, Guildford, UK
11. Colorectal Surgery, Croydon University Hospital, Croydon, UK
12. Radiology, St Mark’s Hospital and Academic Institute, Harrow, UK
13. Colorectal Surgery, Croydon University Hospital, Croydon, UK
14. General Surgery, Wessex Deanery, Winchester, UK
15. Translational Gastroenterology Unit, University of Oxford Nuffield Department of Medicine, Oxford, UK
17. Kleijnen Systematic Reviews Ltd, York, UK
18. Colorectal Surgery, Whips Cross Hospital, Barts Health NHS Trust, London, UK
19. Colorectal Surgery, Nottingham University Hospitals NHS Trust, Nottingham, UK
20. General Practice, Oak Lodge Medical Centre, London, UK
21. Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK
22. University Hospitals of Leicester NHS Trust, Leicester, UK
23. Surgery and Oncology Department, University of Dundee, Dundee, UK

### Correction notice

This article has been corrected since it published Online First. A grammatical error has been corrected.

Twitter: Kevin J Monahan @KevinMonahan, Brian D Nicholson @BrianDNicholson, Richard Booth @richardr2810, Nigel D’Souza @nrmrjeaouza and Robert JC Steele @BobSteele6

Contributors The guideline paper was written and reviewed by all authors, with overview and editing by the first authors. The guideline was developed by the whole group, from inception.

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

Competing interests JEE has served on the clinical advisory boards for Lumendi, Boston Scientific and Paion; has served on the clinical advisory board and has share options in Satisfall Health; and reports speaker fees from Falk, Jannsen and Medtronic. KIM, SB and MM are medical advisors to the charity Bowel Cancer UK. MM is a director of Machesney & Easty Management, a private limited company. JEE is funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR or the Department of Health.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

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

---

REFERENCES


5. Suspected cancer. Recognition and referral (NG12).


Guidelines


