NICE National Institute for Health and Care Excellence

Diagnostics Assessment Programme

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

Committee Papers

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE

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

Diagnostics Assessment Programme

Contents:

1. Overview

2. **External Assessment Report** produced by School of Health and Related Research (ScHARR), The University of Sheffield, the External Assessment Group (EAG) for the assessment
Note, this report is an updated version to the one issued to stakeholders on 1 June 2023. The updates are listed on pages 5 and 6 of the report.

3. Stakeholder comments on the External Assessment Report and EAG responses

4. Addenda to External Assessment Report:
Addendum 1: updated economic analyses
Addendum 2: data used in the statistical syntheses of diagnostic test accuracy

Any information supplied to NICE which has been marked as confidential has been redacted. All personal information has also been redacted.

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

This overview summarises the main issues the diagnostics advisory committee needs to consider. It should be read together with the final scope and the diagnostics assessment report.

1 Aims and scope

Colorectal cancer may be associated with a variety of symptoms including blood in stool (faeces). Several other conditions may present with blood in stools. However, the presence of small amounts of hidden blood in stools (known as faecal occult blood) can indicate that there is bleeding from potentially malignant (cancerous) growths on the inner lining of the large intestine.

Faecal immunochemical testing (FIT) is designed to detect small amounts of blood in a faecal sample by using antibodies specific to human haemoglobin (Hb). A positive FIT alone cannot confirm a diagnosis of colorectal cancer. Further assessment using colonoscopy or alternative testing by CT colonography is required to confirm diagnosis.

Previously FIT had been offered to people with 'low risk' symptoms suggestive of colorectal cancer, while people with 'high risk' symptoms would be immediately referred on the suspected cancer pathway. It has been suggested that FIT should be used for most people with suspected colorectal cancer who present to primary care regardless of risk based on symptoms and age. Clinicians have observed that many people on the suspected colorectal cancer referral pathway do not have any unusual findings at colonoscopy. So, using FIT could mean that people who are unlikely to have colorectal cancer may avoid colonoscopy, and those that are likely to have colorectal cancer can be prioritised more effectively.

There are currently long waiting lists for colonoscopy. In August 2022, 28% of people seen by a specialist for suspected colorectal cancer were not seen within 2 weeks of urgent referral (NHS cancer waiting times, August 2022), although 92% were seen within 28 days. Clinical experts also advised that waiting lists for non-urgent referrals to colonoscopy are currently much longer than the target 18 weeks in some areas of the country. So, using FIT could also release colonoscopy capacity to allow people on non-urgent referral pathways to be seen more quickly.

In 2022, the Association of Coloproctology of Great Britain & Ireland (ACPGBI) and the British Society of Gastroenterology (BSG) published guidance on faecal immunochemical tests in patients with signs or symptoms of suspected colorectal cancer, which recommended that FIT should be used in primary care to prioritise people with clinical features of colorectal cancer for referral for urgent investigation, using a threshold of 10 micrograms Hb per gram of faeces. This guidance was subsequently endorsed by NHS England and implementation has begun in some areas. A full description of the diagnostic pathway in primary care for people with symptoms suggestive of colorectal cancer can be found in section 3.2 of the external assessment report.

Decision question

What is the most clinically and cost-effective way to use quantitative faecal immunochemical tests to reduce the number of people without significant bowel pathology who are referred to the suspected cancer pathway for colorectal cancer, taking into consideration potential colonoscopy capacity constraints for urgent and non-urgent referrals?

Populations

People presenting to primary care with gastrointestinal symptoms or signs indicating a risk of colorectal cancer (excluding people with rectal or anal mass, or anal ulceration).

Specified subgroups for investigation included:

- Age
- Sex
- Ethnicity
- People taking medications or with conditions which increase the risk of gastrointestinal bleeding
- People with blood disorders that could affect the performance of the test (such as beta thalassaemia)
- People with anaemia (including iron deficiency anaemia)

Different threshold values may be needed for these subgroups.

Although FIT is proposed to be offered to the population outlined above, it is possible that introduction of the test would have an indirect impact on people waiting for non-urgent referral to gastroenterology services and/or colonoscopy.

Interventions

Quantitative FIT using specific thresholds of haemoglobin per g of faeces to guide referral. These tests include:

- FOB Gold
- HM-JACKarc
- IDK TurbiFIT
- IDK Hemoglobin (Hb) ELISA
- IDK Hemoglobin/haptoglobin (Hb/Hp) complex ELISA
- NS-Prime
- OC-Sensor iO
- OC-Sensor PLEDIA
- OC-Sensor Ceres
- QuikRead go iFOBT

Comparator

Current practice is standard care according to NICE's guideline on suspected cancer (NG12) and NICE's guidance on quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care (DG30). This includes:

- Clinical assessment and referral for further investigation in secondary care
- Use of FIT (threshold of 10 micrograms Hb/g) to guide referral only for those with 'low risk' symptoms without rectal bleeding (in line with DG30).

Although not considered as a comparator in this assessment, the ACPGBI/BSG guidance has been implemented in some areas of the NHS (see aims).

Healthcare setting

Primary care.

Further details, including descriptions of the interventions, comparator, care pathway and outcomes, are in the final scope.

2 Clinical effectiveness evidence

The external assessment group (EAG) did a systematic review to identify evidence on the clinical effectiveness and diagnostic accuracy of quantitative faecal immunochemical tests (FIT) to detect colorectal cancer, advanced adenomas and inflammatory bowel disease (IBD) at different thresholds. Find the full systematic review results in section 4.3 of the external assessment report.

Overview of included studies

Studies were included if they recruited people presenting to primary care with signs or symptoms indicating a risk of colorectal cancer. Studies using either single FIT (a sample from 1 bowel movement) or dual FIT (a sample from each of 2 different bowel movements) to make referral decisions were included. The EAG categorised the included studies according to how closely the population matched the scope population:

- Type 1: A population that is likely similar to that defined in the scope (see section 1).
- Type 2: A population that is likely representative of the 'high risk' population defined by NICE's guideline on suspected cancer (NG12).
- Type 3: A population that is likely representative of the 'low risk' population defined by NICE's guidance on quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care (DG30).
- Type 4: The population is unclear or unrepresentative of Type 1, 2 or 3 populations. This includes studies where the criteria for FIT or referral were unclear, studies in secondary care, and studies from other countries.

To determine diagnostic accuracy for the different FITs, the EAG considered studies if at least some people had colonoscopy or CT colonography as the reference standard.

Sensitivity analyses were done to examine the effect of including studies in less representative populations, or those in which less than 90% of people had colonoscopy as the main reference standard. Results were similar across these analyses, so the EAG included all study types in its main analysis, which is reported here. For more detail on the sensitivity analyses and sub-population results see sections 4.2 and 4.3 in the external assessment report.

Diagnostic accuracy was determined separately for each FIT manufacturer as equivalence between tests could not be assumed.

Forty-nine studies were included in the review. The evidence base is summarised in Table 1.

Table 1: Number of studies included in the EAG's clinical review.

Device	Single FIT	Dual FIT
FOB Gold	3	-
HM-JACKarc	16	2
IDK TurbiFIT	-	-
IDK Hemoglobin (Hb) ELISA	1	-
IDK Hb/haptoglobin (Hp) complex ELISA	1	-
NS-Prime	1	-
OC-Sensor (all devices)	16	1
QuikRead go	1	1

The EAG assessed study quality using the QUADAS 2 checklist, which assesses the risk of bias and applicability to the decision problem. For the full quality assessment of studies for each test, see section 4.3 and appendix 3 in the external assessment report.

No studies were considered low risk on all items relating to bias. The most common concern was patient selection, either because consecutive samples were not recruited, or because inappropriate exclusions were made, such as excluding people on the basis of not having had a colonoscopy, or not having all blood test results available. The reference standard was also commonly high or unclear risk, either because not all people had colonoscopy or CT colonography, or because it was unclear whether interpretation was blinded to the FIT result. Patient flow scored high risk or unclear in nearly all studies. This was because the time between the FIT and the reference standard was unclear in nearly all studies, people received different reference standards depending on their FIT result or other factors, or people were missing from the study. The index test was generally considered to have a low risk of bias.

The applicability of the population and setting was a common high-risk concern, either because some parts of the target population were excluded (see type 2, 3 or 4 studies above), or because the study population was

unclear. However, the index tests and reference standard target condition were generally at low risk for applicability.

FOB Gold

Three studies were identified that reported diagnostic accuracy data for FOB Gold. Two studies (Benton et al. 2022 and MacLean et al. 2022) used the test with the manufacturer's own SENTiFIT 270 analyser, while the third (Schwettmann et al. 2022) used a Roche device.

Benton et al. 2022 (n=233) reported on a subgroup of the NICE FIT study focused on the NG12 'high risk' population. There was a very small number of colorectal cancer events in the study (n=7; 3.0% prevalence), which can result in less precise estimates of accuracy.

The other 2 studies were considered to have an unclear or unrepresentative population – MacLean et al. (n=553) recruited anyone referred to the 2-week wait pathway in the UK, and Schwettmann et al. (n=163) included people referred to colonoscopy in Norway. The prevalence of colorectal cancer was unusually high in this study (16.0%).

HM-JACKarc

Seventeen studies were identified that reported diagnostic accuracy data for HM-JACKarc, of which 16 were used in the main analysis. Five studies were considered Type 1 (representative population), 3 studies were considered Type 2 ('high risk') and 8 were Type 4 (unclear or unrepresentative). All studies were based in the UK. Sample sizes ranged from 175 to 9,896 participants, and the prevalence of colorectal cancer from 1.1% to 6.4%.

IDK tests

One German study (Sieg et al. 1999) reported data for both the IDK Hb ELISA and the IDK Hb/Hp ELISA. The study was relatively small (n=621) with a low number of colorectal cancer events (n=23; 3.7% prevalence). The population included people with symptoms who were referred to secondary care from

primary care, but specific inclusion criteria were otherwise unclear. Some data was also reported for a combined test (using Hb and Hb/Hp), but the EAG concluded it was not possible to calculate diagnostic test accuracy for this combination, so did not include it in the report.

No studies were identified that reported diagnostic accuracy for the IDK TurbiFIT test. Equivalence data for the TurbiFIT compared to the IDK Hb ELISA was provided by the manufacturer, but the EAG did not consider it to be high enough standard to use in the assessment.

NS-Prime

The EAG identified 1 study (Benton et al. 2022) that reported diagnostic accuracy for the NS-Prime FIT. This was a subgroup of the NICE FIT study that was also used for the FOB Gold analysis and has been described above.

OC-Sensor

Seventeen studies reported diagnostic accuracy data for OC-Sensor, of which 11 were included in the main analysis. Data was available across 4 different devices: iO, PLEDIA, DIANA and MICRO. No data was identified for the CERES device, which is intended to replace the iO. Clinical advice to the EAG was that the devices could be considered equivalent. Sample sizes ranged from 120 to 37,216, and the prevalence of colorectal cancer from 0.6% to 11.7%. Of the studies included in the main analysis, 3 studies were considered Type 1 (representative population), 1 was considered Type 3 (‘low risk’) and 7 were Type 4 (unclear or unrepresentative). Nine studies were in the UK, 1 was in Spain and 1 was in Denmark.

QuikRead Go

One study was identified that reported diagnostic accuracy data on the point-of-care QuikRead go test (MacLean et al. 2021) that recruited people referred from primary care on the 2-week wait pathway (‘high risk’ population). The

study was based in the UK and the prevalence of colorectal cancer was 2.53%.

Dual FIT

Four studies reported data on dual FIT (using 2 samples from different bowel movements before a referral decision is made). Two studies used HM-JACKarc (Gerrard et al. 2023 and Turvill et al. 2018), 1 study used OC-Sensor (Hunt et al. 2022) and 1 used the QuikRead go (Tsapournas et al. 2020). Gerrard 2023 (n=2,637) and Turvill 2018 (n=476) also provided data for single FIT and were included in the main analyses for HM-JACKarc. Tsapournas 2020 (n=242) reported data for single FIT but since the population included people referred from secondary as well as primary care, it was excluded from the main QuikRead go analysis. Hunt 2022 (n=28,622) only reported on dual FIT.

Results

Diagnostic test accuracy for colorectal cancer

Summaries of the diagnostic test accuracy of single FIT for colorectal cancer at selected thresholds are presented in Table 2.

For dual FIT, clinicians stated that they would refer people if either of the 2 FITs were positive (rather than requiring both to be positive to refer). The diagnostic test accuracy results using this strategy are presented in Table 3. Gerrard et al. 2023 and Tsapournas et al. 2020 reported results for both dual and single FIT. In these studies, sensitivity was increased but specificity decreased by using dual FIT.

Summary figures showing sensitivity and specificity for all tests with multiple studies are shown in Figure 1.

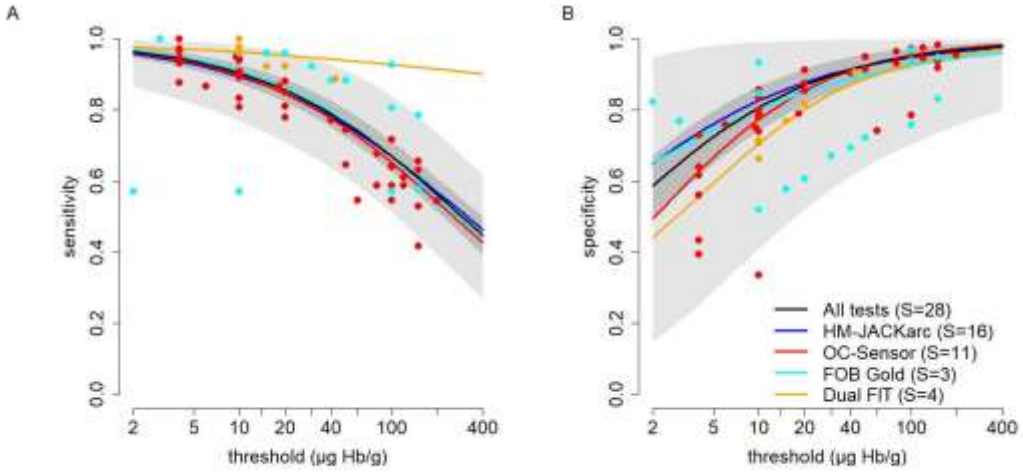


Figure 1: Summary sensitivity (A) and specificity (B) for colorectal cancer for all tests with multiple studies. S, number of studies.

For full test accuracy results see sections 4.3.1 to 4.3.10 of the external assessment report.

Table 2: Summary of single FIT diagnostic accuracy for colorectal cancer at selected thresholds, % (95% CrI or CI)

Threshold (micrograms/g)	FOB Gold (S=3)	HM-JACKarc (S=16)	IDK Hb (S=1)	IDK Hb/Hp (S=1)	NS-Prime (S=1)	OC-Sensor (S=11)	QuikRead go (S=1)
2	Sensitivity: 96.9 (75.6 to 100) Specificity: 65.2 (45.8 to 81.1)	Sensitivity: 95.9 (92.7 to 97.9) Specificity: 65.1 (55.6 to 74.8)	Sensitivity: 87.0 (84.4 to 89.6) Specificity: 88.1 (85.6 to 90.6)	Sensitivity: 82.6 (79.6 to 85.6) Specificity: 80.8 (77.7 to 83.9)	-	-	-
3	Sensitivity: 96.0 (73.9 to 100) Specificity: 69.5 (50.8 to 84.2)	Sensitivity: 94.7 (91.1 to 97.2) Specificity: 70.3 (61.3 to 79.3)	-	-	Sensitivity: 85.7 (48.7 to 97.4) Specificity: 31.9 (26.1 to 38.2)	-	-
4	Sensitivity: 95.1 (72.6 to 100) Specificity: 72.4 (54.3 to 86.2)	Sensitivity: 93.8 (89.8 to 96.5) Specificity: 73.7 (65.1 to 82.2)	-	-	-	Sensitivity: 94.2 (91.2 to 96.7) Specificity: 62.7 (47.4 to 77.2)	-
7	Sensitivity: 93.0 (70 to 99.9) Specificity: 77.5 (60.9 to 89.4)	Sensitivity: 91.4 (86.8 to 94.8) Specificity: 79.6 (71.7 to 87.1)	-	-	-	Sensitivity: 91.8 (88.2 to 94.9) Specificity: 72.3 (58.1 to 84.8)	-
10	Sensitivity: 91.2 (68.2 to 99.8) Specificity: 80.3 (64.9 to 91.1)	Sensitivity: 89.5 (84.6 to 93.4) Specificity: 82.8 (75.2 to 89.6)	-	-	Sensitivity: 71.4 (35.9 to 91.8) Specificity: 83.6 (78.2 to 87.9)	Sensitivity: 89.8 (85.9 to 93.3) Specificity: 77.6 (64.3 to 88.6)	Sensitivity: 92.9 (68.5 to 98.7) Specificity: 70.1 (66.1 to 73.8)

Threshold (micrograms/g)	FOB Gold (S=3)	HM-JACKarc (S=16)	IDK Hb (S=1)	IDK Hb/Hp (S=1)	NS-Prime (S=1)	OC-Sensor (S=11)	QuikRead go (S=1)
20	Sensitivity: 86.4 (64.5 to 99.4) Specificity: 85.1 (71.8 to 93.7)	Sensitivity: 84.7 (79.1 to 89.6) Specificity: 87.9 (81.1 to 93.4)	-	-	-	Sensitivity: 84.7 (80.3 to 89) Specificity: 85.6 (74.5 to 93.6)	-
50	Sensitivity: 76.9 (59.1 to 96.4) Specificity: 89.9 (79.3 to 96.1)	Sensitivity: 75.8 (69.4 to 82.0) Specificity: 92.6 (87 to 96.5)	-	-	-	Sensitivity: 75 (70.2 to 80) Specificity: 92.5 (84.3 to 97.3)	-
100	Sensitivity: 67.0 (53.7 to 88.9) Specificity: 92.6 (83.9 to 97.4)	Sensitivity: 67.0 (60.0 to 74.2) Specificity: 94.9 (90.3 to 97.8)	-	-	Sensitivity: 57.1 (25.1 to 84.2) Specificity: 97.3 (94.3 to 98.8)	Sensitivity: 65.3 (60.2 to 70.7) Specificity: 95.5 (89.4 to 98.6)	Sensitivity: 71.4 (45.4 to 88.3) Specificity: 94.6 (92.4 to 96.2)
150	Sensitivity: 60.2 (48.1 to 81) Specificity: 93.8 (86.2 to 97.9)	Sensitivity: 61.3 (53.7 to 68.9) Specificity: 96.0 (91.9 to 98.4)	-	-	-	Sensitivity: 58.9 (53.4 to 64.7) Specificity: 96.7 (91.6 to 99.1)	Sensitivity: 57.1 (32.6 to 78.6) Specificity: 95.9 (93.9 to 97.3)

CI, confidence interval; CrI, credible interval; S, number of studies.

Table 3: Summary of dual FIT diagnostic test accuracy results for colorectal cancer (either positive), % (95% CI)

Threshold (micro-grams/g)	HM-JACKarc Gerrard 2023	HM-JACKArc Turvill 2018	OC-Sensor Hunt 2022	QuikRead go Tsapournas
10	Sensitivity: 96.6 (90.4 to 99.3) Specificity: 71.2 (69.4 to 73.0)	-	Sensitivity: 98.0 (95.5 to 98.9) Specificity: 66.2 (65.7 to 66.7)	Sensitivity: 100.0 (NR) Specificity: 71.4 (65.5 to 77.3)
15	-	-	-	Sensitivity: 92.3 (77.8 to 100) Specificity: 76.8 (71.3 to 82.3)
20	-	-	-	Sensitivity: 92.3 (77.8 to 100) Specificity: 81.7 (76.6 to 86.8)
43	-	Sensitivity: 87.5 (NR) Specificity: 90.7 (NR)	-	-

CI, confidence interval; NR, not reported.

Comparative diagnostic accuracy between tests

Three studies reported a comparison of 2 or more tests. All 3 studies concluded that different tests would result in different numbers of referrals when using the same threshold. The EAG stated that it is not possible to draw any strong conclusions regarding the comparative performance of the tests, or whether different thresholds would be appropriate for different brands of FIT. This is because there were not enough studies, a low number of colorectal cancer events in 2 of the studies, and no common comparator. For more detail, see section 4.3.11 in the external assessment report.

Diagnostic test accuracy for advanced adenomas and inflammatory bowel disease

Nine studies reported data on FIT accuracy for advanced adenomas and inflammatory bowel disease (IBD). Six studies used the HM-JACKarc, 2 studies used OC-Sensor and 1 study used the QuikRead go.

Results at selected thresholds are presented in Table 4 and Table 5, and a summary in Figure 2. There was large variation between the studies, as well as between tests, so the uncertainty in the accuracy estimates for advanced adenomas and IBD is high. For full details see section 4.3.13 in the external assessment report.

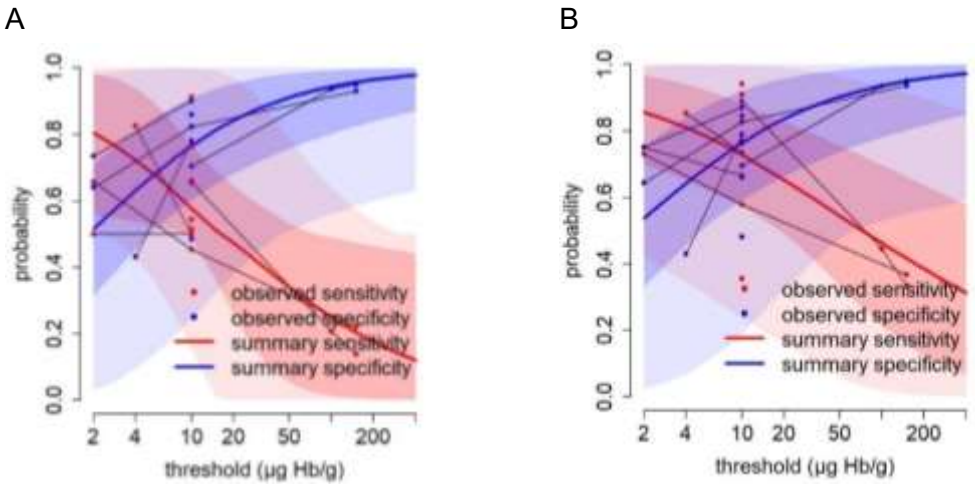


Figure 2: Sensitivity and specificity of FIT (all tests, 9 studies) for advanced adenomas (A) and IBD (B). Dark shading indicates 95% credible interval, light shading indicates 95% predictive interval.

Table 4: Summary of single FIT accuracy for advanced adenomas at selected thresholds, % (95% CrI)

Threshold (micrograms/g)	All tests (S=9)	HM-JACKarc (S=6)	OC-Sensor (S=2)
2	Sensitivity: 80.4 (55.8 to 98.3) Specificity: 51.6 (31.6 to 71.1)	Sensitivity: 59.1 (50.0 to 92.0) Specificity: 55.9 (35.1 to 80.6)	-
3	Sensitivity: 75.9 (54.7 to 96.4) Specificity: 58.4 (39.6 to 77.1)	Sensitivity: 57.7 (50.0 to 89.3) Specificity: 59.9 (41.0 to 84.6)	-
4	Sensitivity: 72.2 (53.7 to 93.8) Specificity: 63.1 (45.6 to 81.0)	Sensitivity: 56.7 (49.8 to 86.8) Specificity: 62.8 (45.1 to 87.3)	Sensitivity: 93.9 (51.5 to 100) Specificity: 46.8 (9.5 to 90.3)
7	Sensitivity: 63.9 (51.4 to 84.6) Specificity: 71.7 (55.3 to 87.7)	Sensitivity: 54.7 (48.2 to 81.2) Specificity: 68.5 (50.8 to 91.7)	Sensitivity: 84.6 (27.8 to 100) Specificity: 70.7 (27.2 to 96.7)
10	Sensitivity: 57.7 (48.6 to 76.7) Specificity: 76.5 (60.3 to 90.9)	Sensitivity: 53.2 (45.9 to 77.6) Specificity: 71.9 (52.0 to 93.8)	Sensitivity: 73.2 (10.1 to 99.9) Specificity: 82.2 (41.6 to 98.7)
20	Sensitivity: 47.4 (26.1 to 64.4) Specificity: 84.2 (68.1 to 95.3)	Sensitivity: 50.9 (37.3 to 71.6) Specificity: 77.9 (53.7 to 96.5)	-
50	Sensitivity: 34.1 (5.6 to 53.2) Specificity: 91.1 (75.7 to 98.2)	Sensitivity: 49.8 (24.3 to 65.7) Specificity: 84.4 (55.4 to 98.5)	-
100	Sensitivity: 25.0 (1.4 to 48.9) Specificity: 94.4 (80.2 to 99.2)	Sensitivity: 48.7 (16.0 to 61.9) Specificity: 88.2 (56.3 to 99.2)	-
150	Sensitivity: 20.4 (0.6 to 47.5) Specificity: 95.7 (82.5 to 99.5)	Sensitivity: 47.8 (12.3 to 60.1) Specificity: 90.1 (56.9 to 99.5)	-

CrI, credible interval; S, number of studies

Table 5: Summary of single FIT accuracy for IBD at selected thresholds, % (95% CrI)

Threshold (micrograms/g)	All tests (S=9)	HM-JACKarc (S=6)	OC-Sensor (S=2)
2	Sensitivity: 85.7 (70.0 to 96.7) Specificity: 53.8 (33.1 to 75.5)	Sensitivity: 86.8 (68.4 to 98.5) Specificity: 57.0 (36.4 to 78.7)	-
3	Sensitivity: 83.1 (67.2 to 95.3) Specificity: 60.0 (40.5 to 80.6)	Sensitivity: 84.9 (66.4 to 97.8) Specificity: 61.3 (41.3 to 82.1)	-
4	Sensitivity: 81.0 (65.1 to 94.0) Specificity: 64.2 (45.8 to 84.0)	Sensitivity: 83.4 (64.7 to 97.1) Specificity: 64.3 (44.8 to 84.4)	Sensitivity: 67.0 (24.7 to 97.9) Specificity: 46.4 (7.4 to 92)
7	Sensitivity: 76.3 (60.4 to 90.7) Specificity: 72.0 (54.7 to 89.4)	Sensitivity: 80.1 (61.2 to 95.3) Specificity: 69.9 (50.8 to 88.1)	Sensitivity: 59.8 (16.4 to 95.5) Specificity: 70.3 (22.3 to 97.5)
10	Sensitivity: 72.9 (57.1 to 88.2) Specificity: 76.4 (59.2 to 92.1)	Sensitivity: 77.6 (58.6 to 94.0) Specificity: 73.3 (53.3 to 90.2)	Sensitivity: 55.1 (12.2 to 93.1) Specificity: 81.9 (35.3 to 99)
20	Sensitivity: 65.3 (49.2 to 82.9) Specificity: 83.6 (66.3 to 95.8)	Sensitivity: 72.3 (52.2 to 91.1) Specificity: 79.2 (57.3 to 93.4)	-
50	Sensitivity: 54.4 (33.6 to 75.5) Specificity: 90.3 (73.7 to 98.3)	Sensitivity: 64.9 (35.1 to 86.9) Specificity: 85.4 (61.7 to 96.2)	-
100	Sensitivity: 46.3 (21.7 to 69.7) Specificity: 93.6 (78.0 to 99.2)	Sensitivity: 59.2 (21.3 to 83.4) Specificity: 89.1 (64.3 to 97.5)	-
150	Sensitivity: 41.7 (15.9 to 66.1) Specificity: 95.0 (80.2 to 99.5)	Sensitivity: 55.7 (15.2 to 81.2) Specificity: 90.8 (65.8 to 98.1)	-

CrI, credible interval; S, number of studies.

Subgroup analyses

Few studies reported characteristics relevant to the subgroups specified in the scope. Where there was little evidence available from the studies included in the main analyses, the EAG relaxed their inclusion criteria to get data on these subgroups. Despite this, the EAG stated that that the evidence base was generally limited and sometimes inconsistent. It was therefore not possible to determine if different FIT thresholds would be appropriate according to these subgroups. For more detail, see section 4.3.12 in the external assessment report.

Anaemia

Eleven studies were identified that reported data on anaemia or iron-deficiency anaemia. Two studies that compared results for people with anaemia to those without reported lower sensitivity and specificity at a threshold of 10 micrograms Hb/g faeces for those with anaemia. One further study reported that the optimal threshold (the point that maximises sensitivity and specificity) for those with anaemia is higher than for those without. In studies that compared people with anaemia to the whole study population, the results were mixed.

Age

Three large studies using HM-JACKarc reported data according to age groups. There was some indication that FIT thresholds may need to be lower in younger people to achieve the same sensitivity as for older people. However, the available data did not provide conclusive evidence that different FIT thresholds should be used or what they should be.

Sex

Three large studies reported data categorised by sex. In 2 studies, a general trend was seen for higher sensitivity and specificity in women compared to men when using thresholds above 10 micrograms/g. This trend was not observed when using thresholds at 10 micrograms/g or below. The third study

concluded that the optimal (highest sensitivity and specificity) threshold was lower for women than for men. However, the available data did not provide conclusive evidence that different FIT thresholds should be used or what they should be.

Medications or conditions that could change the risk of gastrointestinal bleeding

Three studies were found that compared the accuracy of FIT in people who were taking medications that can either increase or reduce the risk of gastrointestinal bleeding. Two studies examined drugs that can increase the risk of gastrointestinal bleeding (antiplatelet, anticoagulant or non-steroidal anti-inflammatory drugs). Both studies observed non-significant reductions in both sensitivity and specificity in people taking the drugs. One study compared FIT in people taking proton-pump inhibitors (PPIs), which may reduce the risk of gastrointestinal bleeding. At a threshold of 20 micrograms/g, sensitivity was similar, and specificity was slightly reduced in people taking PPIs, however no overall conclusion was made on the impact of PPI use on FIT performance for detecting colorectal cancer.

No studies were identified that reported the diagnostic test accuracy of FIT for people with blood disorders (such as beta thalassaemia) that could affect the performance of the test.

Ethnicity

No studies that reported the diagnostic test accuracy of FIT according to ethnicity were identified.

Other outcomes

This section summarises the findings for other outcomes in the scope. Data for these outcomes was only looked for in studies that were included in the EAG's diagnostic test accuracy review. For more detail, please see section 4.3.14 in the external assessment report.

Test failure and test uptake

Test failure rates ranged from 0.2% to 18.8% across 10 studies, with most studies reporting between 2% and 5%. There was no strong evidence that failure rates differed between the tests. Test failure definition was not consistent between studies, so the highest reported value may be due to a particularly broad definition of failure.

Only 2 studies in primary care reported return rate for single FIT. Mowat et al. 2016 had an extremely low return rate (48%), but this may be confounded by the fact that a referral had already been made and so did not depend on the return of the FIT sample. The other study (Bailey et al. 2021), where FIT was being used as part of the diagnostic pathway, reported a return rate of around 91%. In this study, 1% of people who didn't return a FIT were later diagnosed with colorectal cancer.

Three studies of dual FIT reported test return rates. One study found 10.7% returned no FIT, and a further 20.5% returned only 1 FIT. Another reported 4.9% only returned 1 FIT, and the third study had missing samples for 16.1% of participants.

Time-based outcomes

Data on time to reach different points in the diagnostic pathway was reported in 6 studies. The median time to return FIT was 7 days (Cama et al. 2022); the average time to analysis of FIT was 10.1 days (D'Souza et al. 2020); the median time to investigation was 21 days (Gerrard et al. 2023); and the median time to diagnosis was 59 days (Tang et al 2022).

The median time to diagnosis for people with false negative FIT results was 51 days in Cama 2022 and less than 90 days in Bailey 2023. However the diagnostic delay was much longer (more than 1,000 days) for a small group of people, so the interquartile ranges were very large.

Introducing dual FIT increased the median time to investigation by 5 days compared to single FIT in 1 study, while another found that the interval between collection of the 2 samples was 6 days (median).

Diagnostic pathway and referral rates

Mowat et al. (2019 and 2021) reported on the effect of introducing FIT into the diagnostic pathway using a threshold of 10 micrograms/g. Referrals to colorectal services reduced by 9.2% compared to the previous year, and gastroenterology outpatient referrals reduced by 24.1%. Out of 5,372 people tested, 1 person who received a negative FIT result later presented to emergency services.

Patient perspectives

Two studies were found that sought perspectives on FIT from those taking the test using a survey. Both studies included people already referred to the 2-week wait pathway so were considered to be unrepresentative of the scope population.

Both studies addressed the useability of FIT, with 96% and 88% giving a positive response that the device was easy to use. Georgiou Delisle et al. (2022) also found that 90% said the sample was easy to collect.

Around 80% of people were able to overcome feelings of faecal aversion, with a similar proportion saying they preferred FIT to colonoscopy. 96% responded that they would use FIT again.

When asked about satisfaction that a FIT negative result was sufficient to rule out further colonic investigation, 51% responded positively (although respondents to this survey had already had a colonoscopy or CT colonography).

Sociodemographic factors

One study was identified reporting on FIT return rates across demographic subgroups. Significantly higher return rates were observed for women than men, people aged 65 years or older than those below 65, people with a white family background compared to people of other ethnicities (Asian, black or mixed/other ethnic groups analysed separately), and for people in the least socioeconomically deprived quintile compared to those in the most socioeconomically deprived quintile. Strategies to address these disparities could include follow-up after FITs are not returned, providing information in a range of languages, and counselling on the perceived risk of disease and success of treatment.

3 Cost effectiveness evidence

Systematic review of cost-effectiveness evidence

The external assessment group (EAG) did a systematic review to identify any published economic evaluations of faecal immunochemical tests (FIT) in people presenting to primary care with symptoms of colorectal cancer. It searched for any new relevant economic evaluations published since the previous search had been performed during the development of NICE's guidance on quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care (DG30). The aim was to explore the methodological choices in these evaluations, rather than to assess the conclusions reached.

Two economic evaluations (Westwood et al. 2017 and Medina-Lara et al. 2020) were identified that met the EAG's inclusion criteria. These informed the development of the EAG's model. No studies of health-related quality of life were found that met the EAG's inclusion criteria. Find the full systematic review results in section 5.1 of the external assessment report.

Economic analysis

The EAG constructed a de novo economic model to assess the cost effectiveness of different diagnostic strategies using FIT in a primary care setting for people with symptoms of colorectal cancer. The model assessed the health outcomes and costs associated with each strategy over a lifetime horizon, from the perspective of the UK NHS and Personal Social Services. Find a full description of the model in section 5.3 of the external assessment report.

Population

The population was people presenting for the first time to primary care with signs or symptoms suggestive of colorectal cancer. This includes people previously described in NICE's guideline on suspected cancer (NG12) and NICE's guidance on quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care (DG30). People entered the model at an age of 64.9 years. The population was 55% female.

Model structure

The EAG's model consists of a decision tree used to model the diagnostic pathways for a cohort presenting to primary care with symptoms indicating a risk of colorectal cancer, and a state transition model used to model the long-term outcomes and costs. Outcomes were included for people with underlying advanced adenomas and inflammatory bowel disease (IBD) as well as those with colorectal cancer, and those without significant bowel pathology. It was not possible for people in the model to have multiple relevant conditions (so people could not have both underlying cancer and IBD), although people could progress from advanced adenoma to colorectal cancer (see Figure 3).

Three different interventions were compared:

- Intervention 1: All are offered FIT using a single threshold used to define FIT-positive and FIT-negative groups to inform subsequent referral decisions.
- Intervention 2: All are offered FIT using 2 thresholds to define low-, intermediate- and high-risk groups to inform subsequent referral decisions.
- Intervention 3: Referral according to NG12 and DG30. People with ‘high risk’ symptoms are referred directly to the 2-week wait suspected colorectal cancer pathway while people with ‘low risk’ symptoms are offered FIT using a threshold of 10 micrograms/g to inform subsequent referral decisions.

Decision tree

The decision tree part of the model had 3 main branches, reflecting the 3 interventions. Diagrams of these decision trees can be found in Figures 14 to 16 in the external assessment report.

People who completed a FIT and had a positive result (above the single threshold in intervention 1, above the higher threshold in intervention 2, or above 10 micrograms/g in intervention 3) were referred on the 2-week wait pathway, as were people with ‘high risk’ symptoms in intervention 3. Those with a negative result (below the single threshold in intervention 1, below the lower threshold in intervention 2, or below 10 micrograms/g in intervention 3) or those who didn’t complete a FIT were assumed to receive safety netting. Safety netting could consist of:

- referral to the 2-week wait pathway (for people where GPs still have serious clinical concerns despite a negative FIT)
- referral to a non-urgent pathway in secondary care (18-week wait)
- ‘watch and wait’ (management in primary care)
- being offered a repeat FIT.

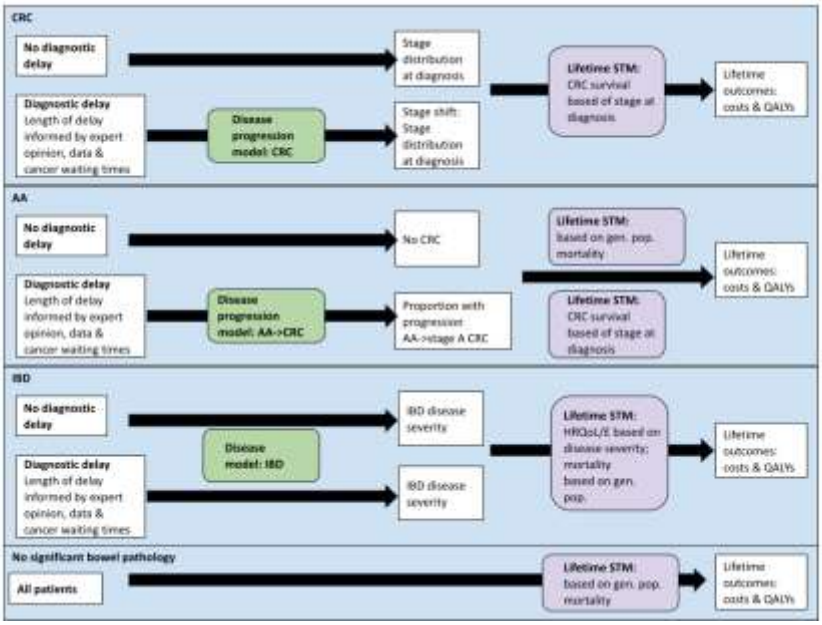
People with intermediate results (between the lower and higher thresholds in intervention 2) received the same options as for those in safety netting, although the proportion of people receiving each option differed (see model inputs).

People referred to secondary care, either through the 2-week or 18-week pathways, received colonoscopy, CT colonography, or other non-invasive investigations. The model assumes that colonoscopy and CT colonography can detect colorectal cancer, advanced adenomas and IBD, but not with perfect sensitivity. People incorrectly diagnosed using these techniques were assumed to be correctly diagnosed by the end of the decision tree, but had a long delay added to the time to diagnosis. People with no significant bowel pathology were assumed to be correctly diagnosed in secondary care (100% specificity).

People managed with 'watch and wait' were assumed to be followed up in primary care and eventually correctly diagnosed either through their GP or through presentation to accident and emergency. For people who were offered a second FIT, the results were not modelled explicitly but a diagnostic delay was added based on some with underlying disease being referred and others being detected after 'watch and wait'.

State transition model

Following the decision tree, a state transition model was used to evaluate long-term outcomes. Different approaches were used for people with different underlying conditions (see Figure 3). The model used a lifetime horizon (up to age 100).



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

Figure 3: Schematic of the EAG's long-term state transition model

For people with underlying colorectal cancer and advanced adenomas, the EAG considered that the length of diagnostic delay could result in disease progression before diagnosis (adenomas developing into cancer, or cancer stage increasing). So, it used a separate model to provide estimates for lifetime health outcomes and costs which incorporated the time to diagnosis. This model assumed that people could only make 1 stage transition in a year. This model is described in Appendix 12 of the external assessment report.

People with IBD or no underlying disease were assumed to enter simple state transition models with 2 states (alive and dead). These used annual cycles in which people had a chance of dying based on general population mortality. The IBD population consisted of people with Crohn's disease and ulcerative colitis. Costs and health-related quality of life for this population were based on the distribution of these conditions and the severity of disease. Diagnostic delay was assumed to increase the probability of complications from IBD for 2 years after diagnosis. For more detail see section 5.3.3 in the external assessment report.

Model inputs

Find the full list of model parameters in section 5.3.4 of the external assessment report.

Some parameters were informed by a survey of the EAG's clinical advisors. Survey results can be found in appendix 11 of the external assessment report.

Test performance

The diagnostic accuracy of each FIT device was obtained from the EAG's evidence synthesis for all thresholds that data was available (described in section 2). FIT return rate was set at 91%, based on Bailey et al. 2021, and test failure was set at 2.1% based on MacDonald et al. 2022.

Safety netting pathways

The EAG modelled 4 options under the safety netting pathway as described in the decision tree structure. The proportion of people who received each option was based on clinical advice to the EAG and is shown in Table 6. The EAG tested 2 scenarios for the intensity of safety netting for people with negative FIT results or who did not return a sample.

Table 6: Proportion of people receiving each form of safety netting after different FIT results

	FIT negative or FIT incomplete (low intensity, base case)	FIT negative or FIT incomplete (high intensity scenario)	Intermediate FIT (Between the 2 thresholds in intervention 2)
Referral to 2WW	5%	15%	85%
Referral to 18WW	10%	25%	10%
Watch and wait	75%	40%	0%
Repeat FIT	10%	20%	5%

2WW, 2-week wait; 18WW, 18-week wait.

In the 'watch and wait' pathway, people with underlying disease were correctly diagnosed after a further 1.9 GP appointments (estimated based on Lyratzopoulos et al. 2013), but 22% of these would only be detected by presentation to accident and emergency (2018 data on routes to cancer

diagnosis, NHS Digital). People diagnosed through this pathway had a significant diagnostic delay.

Investigation in secondary care

People referred to secondary care received colonoscopy (90%), CT colonography (7.5%) or other non-invasive investigations (2.5%). This was based on clinical advice and was intended to reflect availability of CT colonography, as well as preference for non-invasive investigations due to age, frailty or personal choice.

Accuracy of colonoscopy and CT colonography was based on literature (although the specificity of colonoscopy for all conditions was assumed to be 100% due to the nature of the test). Other non-invasive investigations were assumed to have perfect accuracy. For more detail see Table 43 in the external assessment report.

Colonoscopies also had a small associated risk of complications such as bleeding, perforation and death. For more detail see Table 44 in the external assessment report.

Time to diagnosis

Despite the terms ‘2-week wait’ and ’18-week wait’ being used to describe different referral pathways, the actual time from presentation to diagnosis used in the model was based on clinical advice. The EAG also did 2 scenario analyses which used lower (best-case) or higher (worst-case) times to diagnosis to explore this uncertainty. Estimated diagnostic delays are shown in Table 7. These timings do not include the time for initial GP consultation, to complete a FIT and receive the result, or to repeat a failed test, which were all considered negligible.

The EAG assumed that the level of referrals was directly related to waiting times and therefore on time to diagnosis. This is based on the outcomes produced by the model for current practice (intervention 3). For example, if

intervention 1 reduced the number of referrals to the 2-week and 18-week wait pathways by 10% compared to current practice, the EAG assumed that the time to diagnosis on these pathways is also reduced by 10%.

Table 7: Estimated diagnostic delays for people with underlying colorectal cancer, advanced adenomas or IBD in weeks

Pathway	Model base-case	Best-case scenario	Worst-case scenario
2WW, disease diagnosed at referral	2	2	3
18WW, disease diagnosed at referral	27 (6 months)	18 (4 months)	54 (1 year)
2WW or 18WW, disease missed in secondary care	78 (1.5 years)	52 (1 year)	157 (3 years)
Watch and wait	59 (1.1 years)	35 (8 months)	104 (2 years)
Repeat FIT (weighted average of subsequent pathways)	38 (8.7 months)	23 (5.3 months)	69 (1.3 years)

2WW, 2-week wait; 18WW, 18-week wait.

Costs and utilities

Find the full description of costs and utilities used in the model in sections 5.3.4.9 and 5.3.4.10 of the external assessment report.

Test costs

Costs were provided by manufacturers and were generally based on an average cost per test depending on the volume of tests ordered, the analyser used and other factors such as training, servicing and controls. The EAG also included the cost of a GP appointment to discuss the results and take additional blood samples for other tests (cost of blood draw and blood count). The total cost per person for the first FIT completed was between £46.02 and £50.26 depending on the FIT device used. The current practice arm was assumed to use the cheapest possible test. People who did not return the FIT were assumed to incur only the cost of the test (£2.31 for current practice and between £2.31 and £6.46 for interventions 1 and 2). For full detail see Table 48 in the external assessment report.

Costs in secondary care

People referred to secondary care in the model have an appointment with a gastroenterology consultant and 1 of the possible imaging tests (colonoscopy, CT colonography or other non-invasive investigation). People who had a positive result from CT colonography were assumed to also have a confirmatory colonoscopy. People who had colonoscopy were at risk of complications, which had associated costs. Costs and sources are shown in Table 8.

Table 8: Costs for secondary care referral and investigations

Item	Cost	Source (NHS reference costs)
Gastroenterology consultation	£186.48	Face-to-face or non-face-to-face with gastroenterologist
Colonoscopy	£1,003.34	Colonoscopy with biopsy, therapeutic colonoscopy or diagnostic colonoscopy Admission to imaging Follow-up appointment with gastroenterologist
CT colonography	£341.17	CT colonography scan Admission to imaging Follow-up appointment with gastroenterologist
Other non-invasive investigations	£256.29	CT scan (with and without contrast) for 80% of people in this group Follow-up appointment with gastroenterologist for all people in this group
Bleeding complication	£1,695.45	Gastrointestinal bleed procedures with multiple, single or no interventions
Perforation complication	£6,299.74	Major large intestine procedures in adults
Death after perforation	£0.00	Costs of perforation assumed to capture costs before death

Costs of safety netting in primary care

The cost associated with the 'watch and wait' pathway included the cost of an additional 1.9 GP appointments for all people in the pathway, plus the cost of presenting to accident and emergency for 22% of people with underlying disease (see model inputs). The remaining 78% of people with underlying disease were assumed to be eventually referred by the GP to secondary care and incur the costs of consultation and colonoscopy described in Table 8. The

weighted mean cost per person on the ‘watch and wait’ pathway was estimated to be £145.97 (including people without underlying pathology).

People offered a second FIT test under the ‘repeat FIT’ pathway incur the costs of the test, plus 1.9 additional GP appointments. People who didn’t complete the test or had a negative result incur the ‘watch and wait’ costs, while those with positive results from the second test are referred and incur the costs of consultation and colonoscopy described in Table 8. So, the cost per person depends on the specific FIT device and threshold used, but is estimated to be between £339.92 and £586.21 when used at a threshold of 10 micrograms/g.

Lifetime costs and utilities for treating colorectal cancer and adenomas

Baseline lifetime costs and utilities for people with colorectal cancer were generated using the previously published MiMiC Bowel model (Thomas et al. 2020). These outcomes were then adjusted based on additional time to diagnosis and resulting stage shift produced by the EAG’s diagnostic delay model (see appendix 12 in the external assessment report). Table 9 contains a selection of results. People with advanced adenomas were assumed to be asymptomatic if their condition did not progress to colorectal cancer and so have the same costs and quality of life as the general population. Detected adenomas were presumed to be removed at the time of diagnosis.

Table 9: Selected lifetime costs and QALYs resulting from diagnostic delay generated by the EAG's diagnostic delay model

Delay in diagnosis (months)	Costs (advanced adenomas)	QALYs (advanced adenomas)	Costs (colorectal cancer)	QALYs (colorectal cancer)
0	£385	10.35	£11,458	3.37
1	£415	10.34	£11,364	3.33
2	£444	10.33	£11,272	3.29
4	£518	10.31	£11,054	3.19
6	£577	10.29	£10,890	3.13
12	£766	10.23	£10,417	2.94

QALY, quality-adjusted life year

Lifetime costs and utilities for treating IBD

Lifetime costs for people with IBD were taken from Ghosh et al. 2015 and weighted according to disease severity (active, remission or responding to treatment) and type (ulcerative colitis or Crohn's disease). After inflating to 2022 prices, the annual cost of IBD was estimated to be £5,015.75. Utilities for people with IBD were similarly based on disease status and type and were taken from values used in previous NICE assessments. The average utility used in the model for people for IBD was 0.75.

An additional disutility multiplier and annual cost of £399.66 was applied for people who had a long delay to diagnosis of IBD to represent an increased risk of complications in the 2 years after diagnosis. There was no further modelled impact of diagnosis for IBD. For more detail see sections 5.3.4.9 and 5.3.4.10 in the external assessment report.

Base case results

The economic model produced estimates of cost and QALYs for each FIT device across a range of thresholds. Cost-effectiveness estimates were presented as net monetary benefit (NMB) at a willingness-to-pay threshold of £20,000 or £30,000 per QALY. Results presented in this overview use low-intensity safety netting and a willingness-to-pay threshold of £20,000 per

QALY. Find the full results in section 5.3.7 and appendix 14 of the external assessment report.

For intervention 1 (1 threshold), and intervention 2 (2 thresholds), a selection of individual thresholds and threshold pairs were modelled based on clinical input, as there were too many possible combinations to produce useful results.

In general, most implementations of FIT produced a positive NMB compared to current practice, so could be considered cost effective. This was due to cost savings from reducing the number of referrals to secondary care (and so the number of investigations). However, there was also a small reduction in QALYs, because some people got false negative FIT results and so took much longer to get a diagnosis than under current practice. This QALY loss was not completely offset by quicker diagnosis for the majority of people with underlying colorectal cancer resulting from lower demand on colonoscopy services. The overall QALY loss was small across all strategies (around 0.002 QALYs, or less than 1 day of full health for all people in the cohort).

Cost effectiveness of FIT using 1 threshold (intervention 1)

The NMB for FIT when using 1 threshold was positive for all tests at all thresholds except for the NS-Prime at a threshold of 3 micrograms/g (Figure 4). NMB was highest at thresholds around 100 micrograms/g (around £350 per person tested).

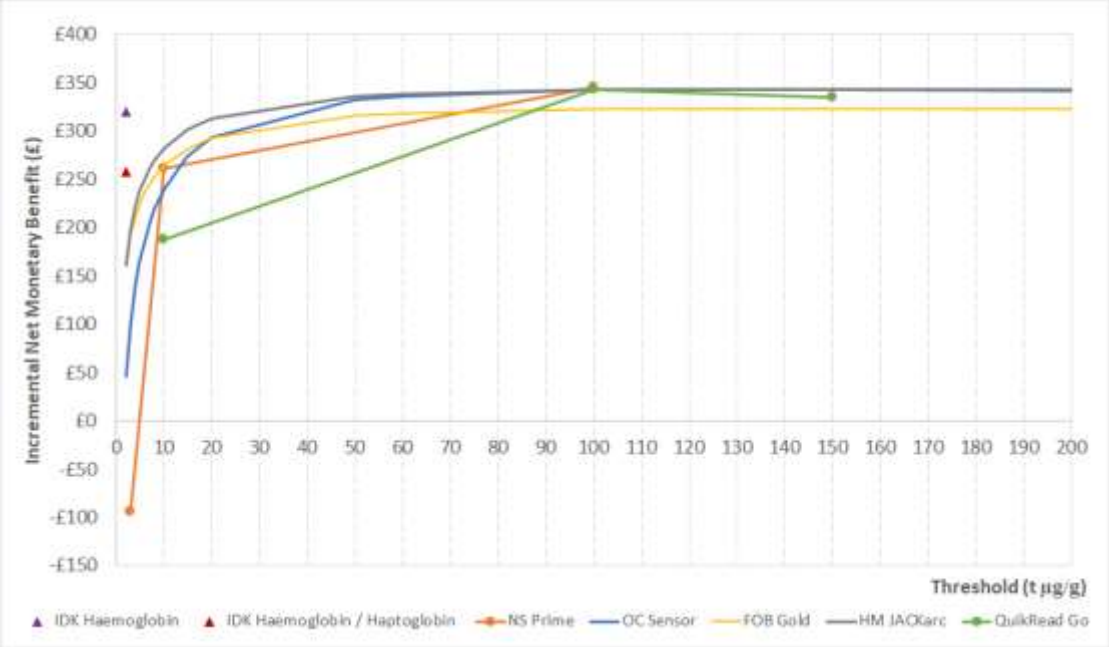


Figure 4: NMB for FIT using 1 threshold, using a willingness-to-pay threshold of £20,000 per QALY and low-intensity safety netting

Impact on secondary care

Using FIT reduced the total number of referrals to secondary care (through 2-week wait or 18-week wait pathways) by around 50% compared to current practice using a single threshold of 10 micrograms/g. This was due to a reduction in the number of 2-week wait referrals, as the total number of 18-week referrals increased.

The reduction in 2-week wait referrals directly translated into a reduction in colonoscopies (Figure 5).

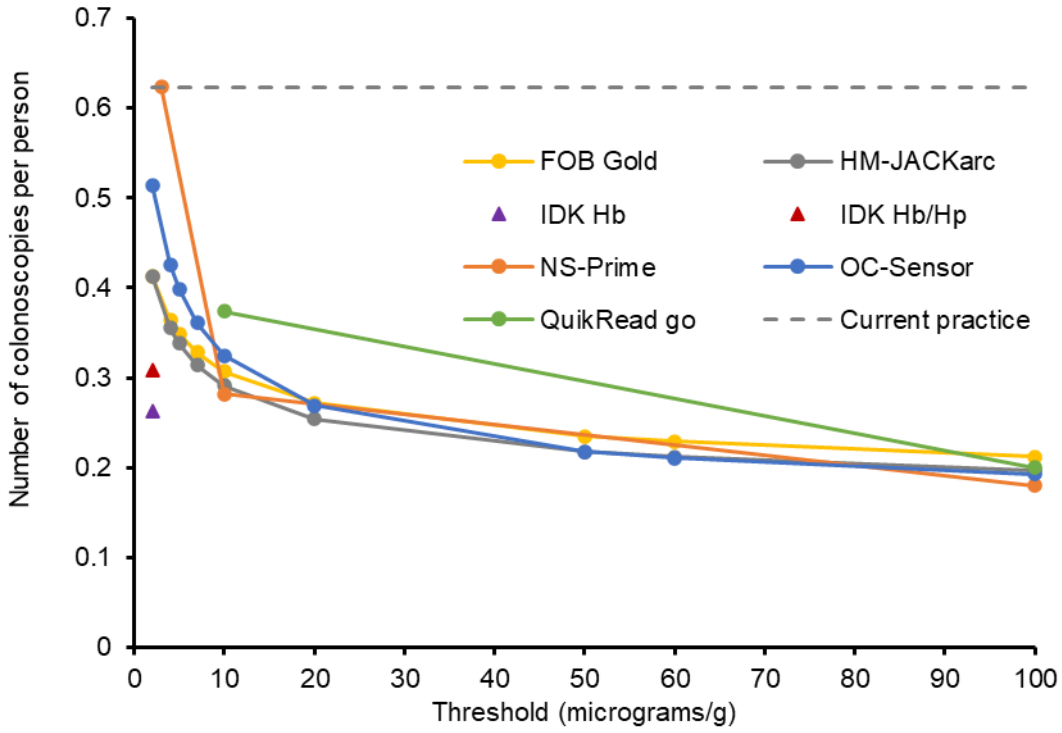


Figure 5: Total number of colonoscopies per person, using FIT with 1 threshold and low-intensity safety netting

Impact on primary care

The number of people managed in primary care correspondingly increased as the number of referrals to secondary care decreased. The number of people managed using 'watch and wait' roughly doubled compared to current practice when using a single threshold of 10 micrograms/g (Figure 6), as did the number of people having repeat FITs.

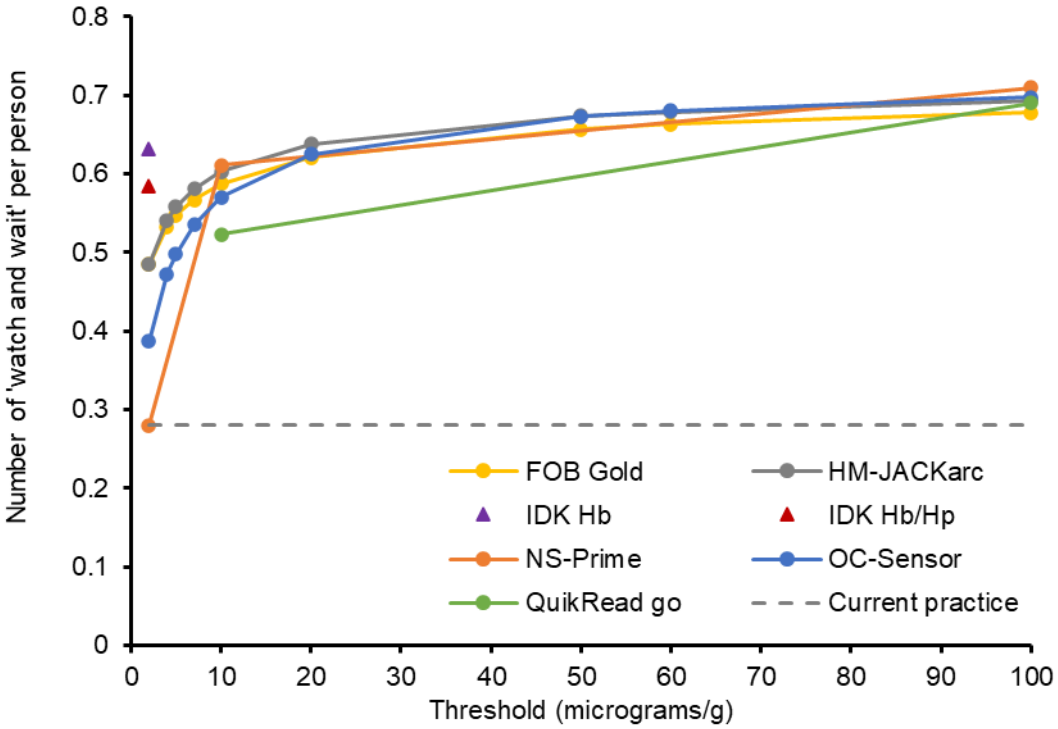
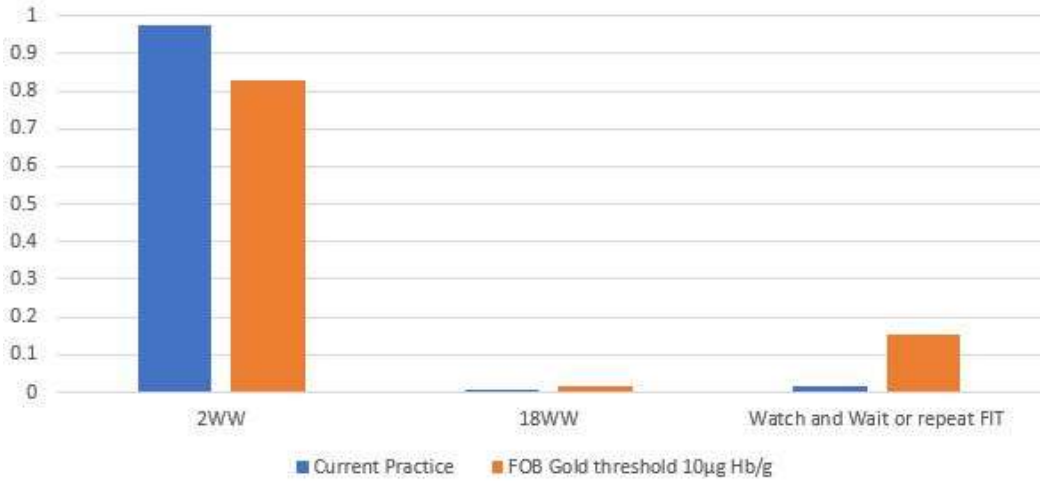


Figure 6: Total number of people managed with 'watch and wait' per person, using FIT with 1 threshold and low-intensity safety netting

The use of FIT also resulted in an increase in the number of people with colorectal cancer who aren't initially referred to secondary care and are diagnosed through watch and wait or repeat FIT. An example is shown in 2WW, 2-week wait; 18WW, 18-week wait

Figure 7, comparing the distribution of pathways that lead to colorectal cancer diagnosis under current practice and when using FOB Gold at a threshold of 10 micrograms/g.



2WW, 2-week wait; 18WW, 18-week wait

Figure 7: Proportion of people with colorectal cancer diagnosed by each pathway when using low-intensity safety netting

Impact on time to diagnosis

Use of FIT increased the average time to diagnosis for people with colorectal cancer, advanced adenomas or IBD. For example, using a threshold of 10 micrograms/g with the FOB Gold test, the average time to diagnosis increased by 1.4 months for people with colorectal cancer, 4.3 months for people with advanced adenomas, and 2 months for people with IBD. This is because some people had a very long delay in diagnosis after getting a false negative FIT result. The time to diagnosis for most people was reduced by a small amount because of the reduced demand for colonoscopy.

FIT using 2 thresholds (intervention 2)

Net monetary benefits for FIT using 2 thresholds were similar to those with 1 threshold, although lower overall, meaning this strategy was estimated to be less cost-effective (Figure 8). All combinations produced positive NMBs except those using the NS-Prime with the lower threshold at 3 micrograms/g. Results for other outcomes were roughly similar to intervention 1 and were most influenced by the value of the lower threshold. For full results using 2 thresholds see section 5.3.7.4 and appendix 14 of the external assessment report.

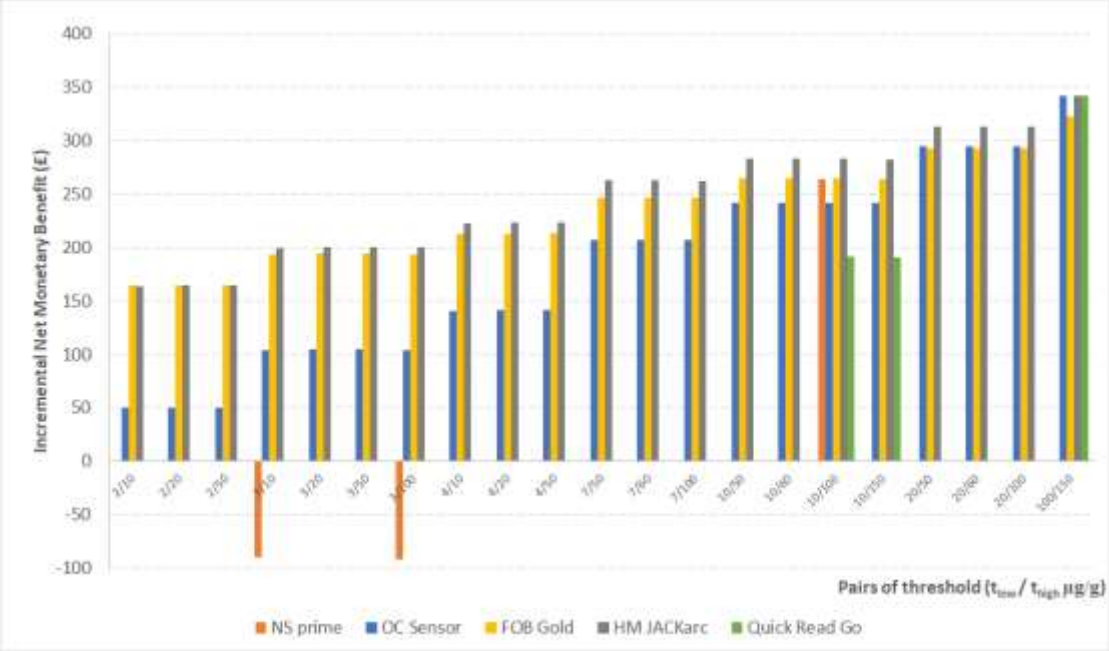


Figure 8: NMB for FIT using 2 thresholds, using a willingness-to-pay threshold of £20,000 per QALY and low-intensity safety netting

Sensitivity analyses

The EAG evaluated uncertainty through probabilistic and deterministic sensitivity analyses. The conclusions of the model were robust to these analyses.

Analysis of alternative scenarios

The EAG ran 11 alternative scenarios to test the findings of the model. As results were broadly similar for all brands of FIT and when using 1 or 2 thresholds, the EAG chose to run the scenarios for HM-JACKarc using a single threshold only. An illustrative summary of the results for a threshold of 10 micrograms/g is presented in Table 10. Find the full list of the EAG's scenario analyses in section 5.3.5.1 of the external assessment report, and the full results in section 5.3.9.

The most cost-effective scenarios were those in which the prevalence of IBD and advanced adenomas was set to 0, and where additional QALYs were lost for increasing diagnostic delay.

The least cost-effective scenario was dual FIT (requiring 2 samples from separate bowel movements). While this increased sensitivity and so reduced the QALY loss, the additional costs from doing twice as many tests were significant (although the NMB was still positive).

Table 10: Scenario analyses for HM-JACKarc using 1 threshold (10 micrograms/g) versus current practice, using low-intensity safety netting

Scenario	Incremental QALYs	Incremental costs	ICER	NMB (£20,000 per QALY)
Base case (deterministic)	−0.0023	−£327	£143,701	£281
Scenario 1: shorter time to diagnosis (best-case, see Table 7)	−0.0012	−£326	£276,014	£303
Scenario 2: longer time to diagnosis (worst-case, see Table 7)	−0.0041	−£327	£78,942	£245
Scenario 3: +1 day QALY loss due to receiving a colonoscopy	−0.0015	−£327	£213,083	£296
Scenario 4: +1 day QALY loss for each month of diagnostic delay	−0.0013	−£327	£249,120	£300
Scenario 5: Dual FIT	−0.0015	−£231	£152,857	£201
Scenario 6: removing IBD and advanced adenomas from the model	−0.0012	−£365	£305,949	£341
Scenario 7: Using FIT return rate of 66% (base-case 91%)	−0.0043	−£362	£84,114	£276
Scenario 8: Use of accuracy data for DG30 low-risk group (intervention 3) from EAG's clinical review analysis for this group	−0.0023	−£289	£125,304	£243
Scenario 9: +1 GP appointment for people with no bowel condition on 'watch and wait' or repeat FIT pathways	−0.0023	−£314	£138,266	£269
Scenario 10: Alternative method to estimate unit costs for FIT in intervention 3 (weighted mean rather than lowest possible cost)	−0.0023	−£327	£143,919	£282
Scenario 11: FIT has perfect accuracy (sensitivity and specificity =1.0) and return rate =1.0	0.0007	−£441	Dominates	£454

IBD, inflammatory bowel disease; ICER, incremental cost-effectiveness ratio; NMB, net monetary benefit; QALY, quality-adjusted life year

4 Summary

Clinical effectiveness

The number of studies identified for the different faecal immunochemical test (FIT) devices was varied, ranging from 1 study for the IDK Hemoglobin (Hb) and IDK Hb/Haptoglobin (Hp) complex ELISAs and the NS-Prime to 17 studies for HM-JACKarc and OC-Sensor. No diagnostic test accuracy data was found for the combined use of IDK Hb + Hb/Hp or for IDK TurbiFIT tests.

Much of the evidence base was in populations that did not directly match the population specified in the scope. This was either because it only included people with ‘high risk’ or ‘low risk’ symptoms (not both), because the population was unclear, or because it was explicitly unrepresentative (for example, including people referred to secondary care). Sensitivity analyses did not find a significant effect on the diagnostic test accuracy results when including these studies, so the external assessment group (EAG) included all study types in its main analysis.

Most studies were considered to have at least some risk of bias. The most common sources of bias were patient selection, reference standard and patient flow. Since a wide range of study types were included in the main analysis, applicability of the population and setting was a common high-risk concern.

Results were reported for each test separately, as there was not enough comparative data to assume that tests could be treated as equivalent. The meta-analysis included data at all reported thresholds and provided summary estimates at all possible thresholds. As an illustrative example, results at a threshold of 10 micrograms/g are shown in Table 11. As the threshold increased sensitivity reduced but specificity increased.

Table 11: Summary diagnostic test accuracy for colorectal cancer at 10 micrograms/g, % (95% CrI or CI)

Test	Sensitivity	Specificity
FOB Gold (S=3)	91.2 (68.2 to 99.8)	80.3 (64.9 to 91.1)
HM-JACKarc (S=16)	89.5 (84.6 to 93.4)	82.8 (75.2 to 89.6)
NS-Prime (S=1)	71.4 (35.9 to 91.8)	83.6 (78.2 to 87.9)
OC-Sensor (S=11)	89.8 (85.9 to 93.3)	77.6 (64.3 to 88.6)
QuikRead go (S=1)	92.9 (68.5 to 98.7)	70.1 (66.1 to 73.8)

CI, confidence interval; CrI, credible interval; S, number of studies.

The study of IDK Hb and IDK Hb/Hp only reported data at 2 micrograms/g. The sensitivity and specificity at this threshold were calculated by the manufacturer to be 87.0% (95% CI 84.4 to 89.6) and 88.1% (95% CI 85.6, 90.6) for the IDK Hb test, and 82.6% (95% CI: 79.6, 85.6) and 80.8% (95% CI: 77.7, 83.9) for the IDK Hb/Hp test.

Four studies reported data on dual FIT. In the studies that reported results for both dual FIT and single FIT, using dual FIT increased sensitivity but decreased specificity.

Data on age, sex and medication use was limited, and no studies reported ethnicity or on blood disorders that could affect FIT results. It was not possible to conclude whether different FIT thresholds would be appropriate for people in these subgroups.

Cost effectiveness

In the EAG's base case, the majority of testing strategies produced a positive net monetary benefit (NMB) compared to current practice, using a willingness-to-pay threshold of £20,000 or £30,000 per quality-adjusted life year (QALY). These findings were consistent across devices and thresholds (whether 1 or 2 thresholds were used). The exception was the NS-Prime where a threshold of

3 micrograms/g was used, which results from the poor estimated specificity of the test from a single study with a low number of colorectal cancer events.

The positive NMB resulted from a cost saving from reducing the number of colonoscopies. However, there was also a small reduction in health from people receiving a false negative FIT result who would previously have had a colonoscopy. The EAG stated that this QALY loss is equivalent to less than 1 day of full health for all people in the cohort.

By reducing the number of colonoscopies, the time to diagnosis for the majority of people was also reduced. However, this was outweighed by the much increased time to diagnosis for people with false negative FIT results, so the overall average time to diagnosis increased compared to current practice.

The EAG's model also predicted that use of FIT would increase the number of people being managed in primary care by a similar amount to the overall reduction in secondary care referrals.

Because of the similarity in NMB results, the simplifications made in modelling, and the uncertainty in many of the model inputs, the EAG stated that it was not possible to clearly identify a specific FIT device and threshold that would be most cost effective. However, the broad conclusion that FIT produces a positive NMB compared to current practice was robust to sensitivity and scenario analyses.

5 Issues for consideration

Clinical effectiveness

Levels of evidence varied across devices from different manufacturers. The most data was available for HM-JACKarc and OC-Sensor, with 17 studies identified for each. Three studies were available for the FOB Gold, but no evidence for the accuracy of IDK TurbiFit or combined IDK Hb + Hb/Hp was

found. Diagnostic test accuracy data for 4 other tests (QuikRead go, NS-Prime, IDK Hb and IDK Hb/Hp) were obtained from a single fairly small study each. The range of thresholds with data available also varied. For the IDK Hb and Hb/Hp test, data was only available at 2 micrograms/g, and for NS-Prime and QuikRead go only 3 thresholds were reported.

Data on comparative diagnostic test accuracy was limited, but all 3 studies identified concluded that at least 1 test produced different results to another test from the same sample. However, because of the limitations of the evidence base, it was not possible to draw clear conclusions on the differences between the tests or whether different thresholds would be appropriate for different tests.

Evidence on population subgroups (age, sex, ethnicity and other factors that could influence the result of a FIT) was limited and in some cases contradictory. No conclusions could be drawn on whether different thresholds would be appropriate for people with these characteristics.

The estimates of diagnostic accuracy of FIT to detect IBD and advanced adenomas were highly uncertain, but were generally lower than the estimates of accuracy to detect colorectal cancer. Introducing FIT may disproportionately affect time to diagnosis and outcomes for people with these conditions.

Cost effectiveness

The EAG stated that the broad conclusion that FIT is cost effective is robust, but this is a result of a cost saving at the expense of a small loss in health. Because of the similarity in results, the simplifications made in modelling, and the uncertainty in many of the model inputs, the EAG stated that it was not possible to clearly identify a specific FIT device and threshold that would be most cost effective. Choice of testing strategy is likely to depend on other factors important to people with gastrointestinal symptoms and healthcare professionals, such as the time to diagnosis or level of referrals.

Results using 2 thresholds followed similar trends to those using a single threshold, but the NMB was generally lower (less cost-effective).

Some key assumptions in the model were informed by expert opinion because of limited data. These included the length of delay in diagnosis for people with false negative results, and how safety netting would be implemented in the NHS for people with negative or intermediate FIT results (see model inputs). This may not be representative of how these people would have their condition managed across the UK.

Part of the rationale for this assessment was that introduction of FIT could reduce the number of referrals for colonoscopy and so reduce waiting times. There was no data to inform this potential effect, so the EAG assumed that the change in number of colonoscopies compared to current practice predicted by the EAG's model had a direct linear impact on the waiting time for colonoscopy.

Dual FIT increased the sensitivity but decreased the specificity compared to using single FIT. It also reduced cost savings as the cost of initial testing included 2 sample kits. This resulted in an overall decrease in NMB (although the QALY loss was also reduced).

6 Equality considerations

NICE is committed to promoting equality of opportunity, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others.

During scoping it was noted that older people and Jewish people of central and eastern European family origin are at increased risk of colorectal cancer. It was identified that the test may not be suitable for people using medicines or with conditions that increase the risk of gastrointestinal bleeding and people with blood disorders, for example sickle beta thalassaemia, in whom faecal

haemoglobin may be difficult to detect. Faecal haemoglobin concentrations may be greater in men than women and may also increase with age.

The EAG looked for evidence on the noted characteristics within studies identified in its review of diagnostic test accuracy. The EAG relaxed their inclusion criteria if the only evidence on these subgroups was in less representative populations (such as those that recruited from both primary and secondary care). However, there was limited evidence in these groups, and the findings were sometimes contradictory. No evidence was found on ethnicity or on blood disorders that could affect the result of a FIT.

Two studies reported that around 90% of people found the FIT tests easy to use (see patient perspectives). However, people with physical or cognitive disabilities may need support to obtain and submit a stool sample using the collection devices, or to understand the purpose of the test and the implications of the test results.

Cultural or demographic preferences may influence the acceptability of tests that require collection of a stool sample. Experience from the bowel cancer screening programme indicates that socioeconomic factors can also act as barriers to engaging with FIT programmes. Evidence from 1 study indicated that age, sex, ethnicity and socioeconomic status all have significant effects on the rate of return of FIT samples (see sociodemographic factors). Strategies to mitigate these differences should be considered. These could include following up after a sample is not returned, providing information in multiple languages, or providing counselling and education services.

7 Implementation

Safety netting

As described in the section on model inputs, the EAG made assumptions about the pathways that would be included in safety netting based on advice from clinical experts. The letter from NHS England that endorsed the guidance

on FIT from the Association of Coloproctologists of Great Britain and Ireland and the British Society of Gastroenterology provided examples of pathways from the North Central London Cancer Alliance and the Oxford University Hospitals NHS trust. However, it is unlikely that these pathways are in place across the UK.

During scoping stakeholders highlighted that in cases where people do not return FIT samples, referral should not be inappropriately delayed. The EAG's model included a pathway for referral for people who did not return samples (as 2-week wait or 18-week wait referral were options in the safety netting arm). This also assumed that GPs would have the option to refer people with negative FIT results if they felt it was necessary.

The EAG's model also predicted that introducing FIT would decrease the number of referrals to secondary care and correspondingly increase the number of people having their condition managed in primary care. Although the secondary care capacity to provide colonoscopy was modelled, the capacity of primary care providers to manage this increased population was not assessed. Clinical advisers suggested that follow-up and monitoring of people in primary care is challenging and clear guidance is needed. Standardisation and improvement of safety netting practice may be required.

Expansion of population

Clinical advisers to the EAG suggested that FIT could be used in a wider spectrum of people than included in the scope for this assessment. It is unclear whether the test accuracy would be similar in a population with less serious or specific symptoms, or what the impact on cost-effectiveness and referral numbers would be. However, this use was out of scope of the assessment.

Reporting of results

During scoping, stakeholders highlighted that different laboratories report FIT results in different ways. Some labs will only report a positive or negative result, others will give the quantitative value for micrograms Hb/g faeces, and some both. Experts said that qualitative reporting can be confusing as if the threshold is not stated then it could easily be confused with the results from a FIT from the screening programme (which uses a different threshold).

8 Authors

Jacob Grant
Topic lead

Judith Shore
Technical adviser

Glossary

Adenoma

A tumour or growth that is not cancerous (benign).

Anaemia

Anaemia is a term to describe conditions where people don't have enough red blood cells or haemoglobin. The most common type is iron-deficiency anaemia, which is caused by low iron levels.

Colonoscopy

An investigation that allows doctors to examine the lining of the colon (large intestine) using a flexible tube that contains a camera and light source (colonoscope).

Computed tomography (CT) colonography

A test that uses CT scans to examine the colon and rectum.

Faecal immunochemical test

A test which detects faecal occult blood using antibodies against human haemoglobin.

Haemoglobin

A protein molecule found in red blood cells. Its presence in faecal samples indicates that gastrointestinal bleeding may be occurring.

Haptoglobin

Haptoglobin is a protein produced by the liver which binds to haemoglobin making it less likely to break down during transit through the gastrointestinal tract. The detection of haptoglobin is claimed to increase the likelihood of detecting lesions higher in the colon.

Inflammatory bowel disease

Inflammatory bowel disease (IBD) is a term used to describe conditions that are characterised by long-term inflammation of the bowel. The 2 main types of IBD are Crohn's disease and ulcerative colitis.

Net monetary benefit

Net monetary benefit (NMB) is a way of describing the cost effectiveness of an intervention at a specific willingness-to-pay threshold (for example, £20,000 per QALY gained). NMB is calculated by multiplying the incremental benefit of the intervention versus current practice by the threshold, and then subtracting the incremental costs.

A positive NMB indicates that the intervention is cost-effective compared to current practice at the given threshold.

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

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

Produced by School of Health and Related Research (ScHARR), The University of Sheffield

Authors Sue Harnan, Senior Research Fellow, Schnarr, University of Sheffield, Sheffield, UK

Aline Navega Biz, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Jean Hamilton, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Sophie Whyte, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Emma Simpson, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Kate (Shijie) Ren, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Katy Cooper, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Mark Clowes, Information Specialist, ScHARR, University of Sheffield, Sheffield, UK

Mr Muti Abulafi, Consultant CR surgeon, Croydon

Alex Ball, Gastroenterologist, Sheffield

Sally Benton, Consultant Biochemist and Clinical Lead for Routine Biochemistry, Royal Surrey NHS Foundation Trust

Richard Booth, Registrar, Croydon

Rachel Carten, Registrar, Croydon

Stephanie Edgar, GP lead for FIT, South Yorkshire CCG

Willie Hamilton, Professor of Primary Care Diagnostics, Exeter

Matt Kurien, Gastroenterologist, Sheffield

Louise Merriman, GP lead for FIT, South Yorkshire CCG

Kevin Monahan, Consultant Gastroenterologist and Endoscopist, London

Laura Heathcote, Research Associate, ScHARR, University of Sheffield, Sheffield, UK

Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK

Correspondence Author Sue Harnan, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK

Date completed Date completed: 21/06/2023

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135637.

Declared competing interests of the authors

None of the authors have any conflicts of interest to declare.

Acknowledgements

We would like to thank Chloe Thomas, Research Fellow, ScHARR, and Olena Mandrik, Research Fellow, ScHARR, for sharing their work (Appendix 12) and advising on the modelling. We would like to thank Paul Tappenden, Professor of Health Technology Assessment, ScHARR, for providing comments on the draft report and Andrea Shippam, Programme Manager, ScHARR, for providing administrative support and in preparing and formatting the report.

Rider on responsibility for report

The views expressed in this report are those of the authors and not necessarily those of the NIHR Evidence Synthesis Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

Harnan S, Navega Biz A, Hamilton J, Whyte S, Simpson E, Ren S, Cooper K, Clowes M, Abulafi M, Ball A, Benton S, Booth R, Carten R, Edgar S, Hamilton W, Kurien M, Merriman L, Monahan K, Heathcote L and Stevenson M. Quantitative faecal immunochemical tests to guide colorectal cancer pathway referral in primary care A Diagnostic Technology Appraisal. School of Health and Related Research (ScHARR), University of Sheffield, 2023.

Contributions of authors

Sue Harnan led the project and was responsible for the day-to-day management. Sue Harnan and Emma Simpson conducted the systematic review of the clinical effectiveness data. Mark Clowes designed and ran the search strategy. Aline Navega Biz, Sophie Whyte and Matt Stevenson conducted the health economic analysis submitted by the company. Jean Hamilton and Kate Ren conducted the statistical aspects of the submission. Muti Abulafi, Alex Ball, Sally Benton, Richard Booth, Rachel Carten, Stephanie Edgar, Willie Hamilton, Matt Kurien, Louise Merriman and Kevin Monahan provided expert clinical opinion to the project. Muti Abulafi, Richard Booth, Rachel Carten and Kevin Monahan also contributed to the systematic review. Laura Heathcote and Sophie Whyte authored Appendix 12. All authors were involved in drafting and commenting on the final report.

Standard copyright statement on the front page of the ERG/AG report:

About ScHARR

The School of Health and Related Research (ScHARR) is one of the four Schools that comprise the Faculty of Medicine at the University of Sheffield. ScHARR brings together a wide range of medical- and health-related disciplines including public health, general practice, mental health, epidemiology, health economics, management sciences, medical statistics, operational research and information science. It includes the Sheffield unit of the Trent Institute for Health Services Research, which is funded by NHS R&D to facilitate high-quality health services research and capacity development.

The ScHARR Technology Assessment Group (ScHARR-TAG) synthesises research on the clinical effectiveness and cost-effectiveness of healthcare interventions for the NIHR Health Technology Assessment Programme on behalf of a range of policy makers, including the National Institute for Health and Care Excellence (NICE). ScHARR-TAG is part of a wider collaboration of a number of units from other regions including Southampton Health Technology Assessment Centre (SHTAC), University of Southampton; Aberdeen Health Technology Assessment Group (Aberdeen HTA Group), University of Aberdeen; Liverpool Reviews & Implementation Group (LRiG), University of Liverpool; Peninsular Technology Assessment Group (PenTAG), University of Exeter; the NHS Centre for Reviews and Dissemination, University of York; Warwick Evidence, The University of Warwick; the BMJ Group and Kleijnen Systematic Reviews.

No CIC (Commercial in Confidence) or AIC (academic in confidence) data is included in this report.

Changes to original report

The original EAG report was dated 24th May 2023.

Edits were made in response to Stakeholder comments, and a revised report dated 21st June 2023 prepared. The edits are described in this Table.

In addition, an addendum was prepared that updated modelling results using a new price for HM-JACKarc.

Location in report	**Edit made**
Cover sheet	Date of report amended to resubmission date.
	Statement relating to ACIC data changed as the report contains no ACIC data.
Executive summary section 2.4.1	Information relating to OC-Sensor Ceres and equivalence data added
	Typo corrected in: "In studies that reported estimates for both, the sensitivity was higher and specificity was lower when using Dual FIT (test interpreted as positive if either FIT positive) than that achieved when using only the first FIT test result to interpret the test.
Executive summary section 2.5	Information relating to equivalence data for IDK tests and OC-Sensor Ceres added
Main report, throughout	Table and figure numbers have been updated, and some references added, affecting reference numbers throughout.
Main report section 3.3.3.2	Changed name of manufacturer for HM-JACKarc
Main report section 3.3.3.4	CIC removed from OC-Sensor Ceres information.
Main report Table 1	Limit of detection and quantification changed for HM-JACKarc
Main report section 3.3.8.9	Edit relating to data from DG30.
Main report, section 4.1.9	Underline removed from QUADAS-2
Main report, Section 4.3.2	Typo correct in "Sample sizes ranged from 120[81] to 37,216[32] and CRC prevalence from 0.59%[45] to 11.65%.[73]"
	Paragraph added on equivalence data for OC-Ceres.
Main report, Section 4.3.6	Paragraph on equivalence data for IDK TurbiFIT and ELISA tests edited
	Section headed "IDK Hb and Hb/Hp tests": added "and the EAG could not validate this statement".
Main report, Section 4.3.10	Text relating to interpretation of 95% CIs edited.

	Figure 10 updated to remove Dual FIT data.
	Table 19 Several small edits/corrections made, including to footnote b.
Main report, Section 4.3.14.1	Edit to sentence "Data were available for OC-Sensor and HM-JACKarc only; there was no strong evidence that rates differed according to these test brands."
	In "Uptake" section, added "All studies took place in secondary care."
Main report, Section 5.3.4.10	Text relating to variable costs added in section titled "FIT costs"
	Table 48, text relating to costs edited in column 3, OC-Sensor
Main report, Table 55-58, Table 60, 61	Data in row "Reduction in number of COLs" corrected.
Main report, Table 67	Data in rows "Reduction in number of referrals (2WW only); Increase in number of referrals (18WW only); Reduction in number of COLs" corrected.
Main report, Section 6.1.1	Typo corrected in section titled "Dual FIT".
	"; all Dual FIT studies were in secondary care." added in section "Test failures, uptake and repeat tests"
Main report, Section 6.2.1.1	Edits made relating to equivalence data for OC-Sensor Ceres and IDK tests.
	Edit made to note test failure, uptake and repeat tests data largely from HM-JACKarc and OC-Sensor
Main report, Section 6.2.1.1	Capitalisation removed from "Limitations".
Main report, Section 6.2.1.2	Limitation added relating to the impact of distribution and sample return methods and causes of test failures.
Main report, Section 6.4	Text added relating to test failures, uptake and repeat tests and Dual FIT.
Main report, Section 6.5	Text added relating to FIT distribution and avoiding test failures.
Appendix 5	Data on NPV and PPV added

Table of contents

1 LIST OF ABBREVIATIONS 1
2 EXECUTIVE SUMMARY 3
2.1 Background 3
2.2 Objectives 3
2.3 Methods 4
2.4 Results 6
2.5 Discussion 9
2.6 Implications for service provision 9
2.7 Suggested research priorities 9
3 BACKGROUND AND DEFINITION OF THE DECISION PROBLEM 11
3.1 Condition and aetiology 11
3.2 Current service provision 12
3.3 Description of the decision problem 17
4 CLINICAL EVIDENCE 33
4.1 Methods 33
4.2 The analysis plan and rationale 41
4.3 Results 45
4.4 Data selected to enter the cost-effectiveness model 146
5 COST-EFFECTIVENESS 149
5.1. Review of existing health economic analyses published 149
5.2. Review and critical appraisal of economic analyses provided by test manufacturers 159
5.3. Independent economic evaluation 159
6 DISCUSSION AND CONCLUSIONS 219
6.1 Statement of principal findings 219
6.2 Strengths and limitations of the assessment 224
6.3 Uncertainties 227
6.4 Generalisability 228
6.5 Implications for service provision 228
6.6 Suggested research priorities 228
6.7 The use of patient and public involvement 229
6.8 Equality, Diversity, and Inclusion 229
7 REFERENCES 230
8 APPENDICES 247
Appendix 1: Literature search strategies 247
Appendix 2: Conversion of sensitivity and specificity to TP, TN, FP, FN 258

Appendix 3: QUADAS scoring scheme and scores with reasons for all studies 259
Appendix 4: Statistical methods for the evidence synthesis..... 294
Appendix 5: Additional meta-analysis results and NPV and PPV results..... 296
Appendix 6: Worked example relating to potential bias of excluding FIT negative patients 298
Appendix 7: Clinical review: Table of excluded studies with rationale..... 302
Appendix 8: Diagnostic test accuracy data entering the statistical synthesis..... 303
Appendix 9: Test uptake and repeat tests data from secondary care..... 304
Appendix 10: Review of economic evaluations and HRQoL studies: table of excluded studies with rationale and quality assessment of studies included..... 306
Appendix 11: EAG survey collected from EAG's clinical advisors..... 311
Appendix 12: Model estimation of the impact of additional time to diagnosis on colorectal cancer outcomes 325
Appendix 13: Methods for pooling prevalence data from the EAG clinical review.... 336
Appendix 14: Additional health economic results 337

Tables

Table 1: Summary of interventions (adapted from Table 1 from the NICE scope[21]) 25
Table 2: Summary of studies entering the analysis, by test and study population..... 45
Table 3: Study and patient characteristics of HM-JACKarc studies 51
Table 4: HM-JACKarc studies: EAG's assessment of risk of bias and applicability 62
Table 5: Summary sensitivity and specificity at specific thresholds for HM-JACKarc 66
Table 6 : Study and patient characteristics of OC-Sensor studies..... 70
Table 7: OC-Sensor studies: ScHARR's assessment of risk of bias and applicability 77
Table 8: Summary sensitivity and specificity at specific thresholds for OC-Sensor .. 79
Table 9: Study and patient characteristics of FOB-Gold studies..... 82
Table 10: FOB-Gold studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores..... 84
Table 11: Study and patient characteristics and diagnostic test accuracy of QuikRead go studies..... 87
Table 12: QuikRead go studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores..... 88
Table 13: Study and patient characteristics and diagnostic test accuracy of the NS-Prime study 90
Table 14: NS-Prime study: ScHARR's assessment of risk of bias and applicability 90
Table 15: IDK studies: ScHARR's assessment of risk of bias and applicability 93
Table 16: Study and patient characteristics of IDK Hemoglobin (human) and hemoglobin/haptoglobin complex ELISA tests..... 94
Table 17: Study and patient characteristics of studies reporting data for dual FIT 96
Table 18: Summary sensitivity and specificity at specific thresholds for all tests together 99

Table 19: Summary sensitivity and specificity at selected thresholds 105
Table 20: Sensitivity and specificity reported in studies comparing different tests within the same patients..... 108
Table 21: Sensitivity and specificity of studies reporting data for patients with anaemia 111
Table 22: Sensitivity and specificity by age..... 115
Table 23: Sensitivity and specificity by sex 118
Table 24: Sensitivity and specificity for patients taking medications that may affect the risk of GI bleeding..... 121
Table 25: Studies reporting data on AA and IBD 124
Table 26: Summary sensitivity and specificity at selected thresholds for AA outcome 127
Table 27: Summary sensitivity and specificity at selected thresholds for IBD outcomes 129
Table 28: Studies issuing FIT in primary care or issuing DUAL FIT, and reporting test failure rates, test uptake and number of repeat tests..... 131
Table 29: Studies reporting other outcomes listed in the NICE scope 135
Table 30: Study characteristics of patient acceptability studies..... 139
Table 31: Patient characteristics of patient acceptability studies 140
Table 32: Patient acceptability results, Georgiou Delisle study[34]..... 142
Table 33: Table Patient acceptability results, MacLean study..... 142
Table 34: Study characteristics of equity study[33] 145
Table 35: Results of equity study Bailey et al. 2023[33] 147
Table 36: Existing economic evaluations – analytic scope..... 154
Table *37*: Existing economic evaluations - modelling approach, main assumptions, definition of health states and summary of HRQoL included..... 155
Table 38: Scope of the EAG economic analysis 159
Table 39: Evidence sources used in the model..... 170
Table 40: Parameters related to disease prevalence and disease stage or severity used in the base-case analysis..... 172
Table *41*: Proportion of patients receiving each of the management pathways following FIT results*, based on clinical advice provided to the EAG..... 175
Table 42: Proportion of patients receiving each pathway at their lower GI referral (2WW/18WW) 176
Table *43*: Estimates of accuracy for imaging tests used in patients in 2WW and 18WW 177
Table 44: Complications and QALY losses associated with colonoscopy included in the short-term model..... 178
Table 45: Estimated diagnostic delays by each type of pathway and diagnostic result† 179
Table 46: Health utilities applied in the EAG model for IBD patients..... 180
Table 47: Proportion of patients by disease severity and type applied in the EAG model for IBD patients..... 181
Table 48: Test costs assumed in EAG analysis 181
Table 49: Costs related to complications of colonoscopy..... 184
Table 50: Distributions used in EAG probabilistic analyses 185
Table 51: Accuracy estimates used in Scenario 8 for detection of CRC..... 187
Table 52: Tabulated results for FOB Gold using one threshold 194

Table 53: Tabulated results for FOB Gold using two thresholds 195
Table 54: Tabulated results for HM JACKarc using one threshold 196
Table 55: Tabulated results for HM JACKarc using two thresholds 197
Table 56: Tabulated results for IDK Haemoglobin using one threshold 199
Table 57: Tabulated results for IDK Haemoglobin using one threshold 200
Table 58: Tabulated results for NS Prime 201
Table 59: Tabulated results for OC-Sensor using one threshold 202
Table 60: Tabulated results for OC-Sensor using two thresholds 203
Table 61: Tabulated results for QuikRead go 205
Table 62: Deterministic sensitivity analyses results for HM JACKarc using one threshold (10 µg/g) 207
Table 63: Tabulated results for HM JACKarc using one threshold 208
Table 64: Tabulated results for HM JACKarc using one threshold 209
Table 65: Tabulated results for HM JACKarc using one threshold 210
Table 66: Tabulated results for HM JACKarc using one threshold 211
Table 67: Tabulated results for HM JACKarc using one threshold 212
Table 68: Tabulated results for HM JACKarc using one threshold 213
Table 69: Tabulated results for HM JACKarc using one threshold 214
Table 70: Tabulated results for HM JACKarc using one threshold 215
Table 71: Tabulated results for HM JACKarc using one threshold 216
Table 72: Tabulated results for HM JACKarc using one threshold 217
Table 73: Tabulated results for HM JACKarc using one threshold 218
Table 74: HM-JACKarc studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores 263
Table 75: OC-Sensor studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores 274
Table 76: FOB-Gold studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores 285
Table 77: QuikRead go studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores 288
Table 78: NS-Prime study: ScHARR's assessment of risk of bias and applicability, with reasons for scores 290
Table 79: FOB-Gold studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores 292
Table 80: Parameters used to inform priors for syntheses with less than 5 studies. 295
Table 81: Meta-analysis model fit statistics 296
Table 82: PPV and NPV for selected thresholds, from the synthesised sensitivity and specificity, or individual studies where no synthesis was performed 297
Table 83: Proportion of patients receiving each of the management pathways following FIT results, clinicians individual answers 321
Table 84: Proportion of patients receiving each imaging test as the first test during 2WW referral 322
Table 85: Proportion of patients receiving each imaging test as the first test during 18WW referral 323
Table 86: Proportion of patients receiving each of the management pathways following FIT results, estimated values by the EAG based on clinicians' answers 324

Table 87: Proportion of patients receiving each main imaging test during their lower GI referral (2WW/18WW) 324
Table 88: Time to diagnosis by type of pathway and diagnostic result (weeks), based on clincians ansewers to questionaire 324
Table 89: age distributions 328
Table 90: transition probabilities used within the model [96] 329
Table 91: expected discounted LYs, QALYs, and inflated treatment costs by age and stage at diagnosis 331
Table 92: estimated outcomes by additional time to diagnosis for CRC (NMB = Net Monetary Benefit, WTP = Willingness To Pay). All outcomes are discounted at 3.5% per annum. 334
Table 93: impact of additional time to diagnosis of HRA (NMB = Net Monetary Benefit, WTP = Willingness To Pay). All outcomes are discounted at 3.5% per annum. 335
Table 94: Tabulated results for FOB Gold using one threshold 340
Table 95: Tabulated results for FOB Gold using two thresholds 341
Table 96: Tabulated results for HM-JACKarc using one threshold 343
Table 97: Tabulated results for HM-JACKarc using two thresholds 344
Table 98: Tabulated results for IDK Haemoglobin using one threshold 345
Table 99: Tabulated results for IDK Haemoglobin/Hapto using one threshold 346
Table 100: Tabulated results for NS Prime 347
Table 101: Tabulated results for OC Sensor using one threshold 348
Table 102: Tabulated results for OCSensor using two thresholds 349
Table 103: Tabulated results for QuikRead go 350

Figures

Figure 1: The diagnostic pathway for patients presenting to primary care with symptoms of colorectal cancer. Based on NG12[10] and DG30[11] 14
Figure 2 The diagnostic pathway for patients presenting to primary care with symptoms of colorectal cancer as recommended in the ACPGBI/BSG guideline[12] 16
Figure 3: Proposed new pathway incorporating FIT for all patients in primary care: A) using a single FIT threshold; and B) using two FIT thresholds to create an intermediate risk group who would follow a different diagnostic pathway 20
Figure 4: Prisma flow diagram of study selection 47
Figure 5: Observed data and summary sensitivity and specificity for HM JACKarc. 65
Figure 6: Observed data and summary sensitivity and specificity for OC-Sensor 79
Figure 7: Observed data and summary sensitivity and specificity for FOB Gold 85
Figure 8: Observed data and summary sensitivity and specificity for all tests together 98
Figure 9: Summary sensitivity and specificity, reference standard sensitivity analysis 103
Figure 10: Sensitivity and specificity for all tests 104
Figure *11*: Observed data and summary sensitivity and specificity for AA outcomes. A) All tests, ROC B) All tests as a function of threshold, C) HM-JACKarc D) OC-Sensor 126
Figure 12: Observed data and summary sensitivity and specificity for All tests. IBD outcomes 128

Figure *13*: PRISMA flow diagram, review for economic evaluations and HRQoL studies 153
Figure 14: EAG model - decision tree structure, Intervention 1 (FIT with one threshold of t µg/g) 161
Figure 15: EAG model - decision tree structure, Intervention 2 (FIT with two thresholds of t_{high} and t_{low} µg/g)..... 162
Figure 16: EAG model - decision tree structure, Intervention 3 (DG30 and NG12 recommendations, FIT for DG30 low-risk patients with threshold of 10 µg/g)..... 162
Figure 17: EAG model - state transition model..... 163
Figure 18: NMB for Intervention 1 assuming a threshold of £20,000 per QALY gained and low intensity safety netting..... 189
Figure 19: NMB for Intervention 2 assuming a threshold of £20,000 per QALY gained and low intensity safety netting..... 191
Figure 20: The proportion of patients diagnosed by category 192
Figure 21: model diagram..... 327
Figure 22: stage distribution of colorectal cancer in the 2WW population in England [281] 328
Figure 23: change in CRC stage distribution by additional time to diagnosis 331
Figure 24: iNMB for Intervention 1 assuming a threshold of £20,000 per QALY gained and high intensity safety netting 337
Figure 25: iNMB for Intervention 2 assuming a threshold of £20,000 per QALY gained and high intensity safety netting 337
Figure 26: iNMB for Intervention 1 assuming a threshold of £30,000 per QALY gained and low intensity safety netting..... 338
Figure 27: iNMB for Intervention 2 assuming a threshold of £30,000 per QALY gained and low intensity safety netting..... 338
Figure 28: iNMB for Intervention 1 assuming a threshold of £30,000 per QALY gained and high intensity safety netting 339
Figure 29: iNMB for Intervention 2 assuming a threshold of £30,000 per QALY gained and high intensity safety netting 339

1 LIST OF ABBREVIATIONS

µg/g or µg Hb/g	Micrograms of haemoglobin per gramme of faeces
18WW	18 week wait
2WW	Two week wait
A&E	Accident and emergency
AAs	Advanced adenomas
ACPGBI	The Association of Coloproctology of Great Britain and Ireland
AEs	Adverse events
BSG	British Society of Gastroenterology
CCTs	Controlled clinical trials
CD	Crohn's disease
CG27	Clinical guideline 27
CI	Confidence interval
COL	Colonoscopy
CRC	Colorectal cancer
CRD	Centre for Reviews and Dissemination
CrI	Credible intervals
CT	Computerised tomography
CTC	Computed tomography colonography
DALY	Disability-adjusted life year
DAP	Diagnostics assessment programme
DG30	Diagnostics guidance 30
DIC	Deviance information criterion
DSA	Deterministic sensitivity analyses
EAG	Evidence assessment group
ELISA	Enzyme-linked immunosorbent assay
eRS	Electronic Referral System
FIT	Faecal immunochemical test
FNs	False negatives
fOBT	Faecal occult blood test
FPs	False positives
GI	Gastrointestinal
GP	General practitioners'
HRQoL	Health-related quality of life

IBD	Inflammatory bowel disease
ICERs	Incremental cost-effectiveness ratios
iFOBT	Immunochemical faecal occult blood test
iNMB	Incremental net monetary benefit
LYG	Life years gained
NCRAS	National Cancer Registration and Analysis Service
NG12	National guideline 12
NHB	Net health benefit
NHS	National health service
NMB	Net monetary benefit
NSBP	No significant bowel pathology
OGD	Oesophago-gastroduodenoscopy
PrI	Prediction intervals
PROSPERO	International prospective register of systematic reviews
PSA	Probabilistic sensitivity analysis
PSS	Personal social services
QALYs	Quality-adjusted life years
QUADAS-2	Quality assessment of diagnostic test accuracy studies version 2
RCT	Randomised control trial
ROC	Receiver operating characteristic
SNOUT	Sensitive test when Negative rules OUT the disease
TNs	True negatives
TPs	True positives
UC	Ulcerative colitis

2 EXECUTIVE SUMMARY

2.1 Background

Early diagnosis and treatment of colorectal cancer (CRC) in people presenting to primary care with symptoms can improve survival and cure rates. The introduction of the NICE National Guideline 12 (NG12) guidelines in 2015 to expand symptoms-based criteria for referral to secondary care led to an increase in the number of urgent two week wait (2WW) referrals, but no corresponding increase in the proportion of patients investigated through the 2WW who had cancer. This has led to pressure on colonoscopy capacity and to long waiting times in some areas, especially in the non-urgent (18 week wait (18 WW)) referral pathway.

Quantitative faecal immunochemical tests (FIT) are designed to detect occult (small amounts) of blood in stool samples (faecal haemoglobin) using antibodies specific to human haemoglobin. They are currently used in patients with low-risk symptoms in primary care (as described in Diagnostics guidance 30 (DG30)), but not in patients with high/medium-risk symptoms as defined in NG12, who are instead referred directly to secondary care. There is evidence that FIT is a better predictor of CRC risk in patients than symptoms alone and could result in fewer referrals of people without CRC to the urgent 2WW suspected CRC pathway. Therefore, triage with FIT for all patients could mean that people who are unlikely to have CRC may avoid unnecessary referral for investigations. They would also avoid the associated disadvantages of referral such as patient anxiety, time off work and loss of economic productivity, as well as the rare adverse events associated with colonoscopy for example, bleeding, perforation and death. Furthermore, those that are likely to have CRC could be prioritised more effectively, potentially leading to a reduction in time to diagnosis. This may also release colonoscopy capacity to allow people on non-urgent referral pathways to be seen more quickly. The extent to which colonoscopy capacity is released and time to diagnosis affected will be dependent in part on the threshold used to define a positive FIT, with a higher threshold resulting in fewer referrals, but a greater chance of missing disease.

2.2 Objectives

The decision problem in the NICE scope was "*What is the most clinically and cost-effective way to use quantitative faecal immunochemical tests to reduce the number of people without significant bowel pathology who are referred to the suspected cancer pathway for colorectal cancer, taking into consideration potential colonoscopy capacity constraints for urgent and non-urgent referrals?*" Eight FIT tests were within the scope of the assessment, namely HM-JACKarc, FOB Gold, OC-Sensor, NS Prime, IDK TurbiFIT, IDK Hemoglobin ELISA (IDK Hb), IDK Hb/Hp complex ELISA (IDK Hb/Hp), and QuikRead go.

The decision problem was addressed through a systematic review of evidence relating to the accuracy of the tests, a statistical synthesis to pool data across studies and an economic model which aimed to estimate the cost effectiveness of FIT strategies based on diagnostic accuracy, the number of colonoscopies undertaken and the impact on time to diagnosis.

2.3 Methods

2.3.1 Clinical evidence review methods

Searches were conducted across four databases and six registries in December 2022. The titles and abstracts of records retrieved were screened by one reviewer, with the first 20% checked by a second reviewer before the remainder were screened. Records for which the full text were obtained were checked for inclusion by a second reviewer. Data extraction and quality assessment were conducted by one reviewer and checked by a second. Study quality was assessed using QUADAS-2.

Since no randomised controlled trials were identified in any of the tests, studies were included if they reported the diagnostic test accuracy of FIT in patients presenting to primary care with signs or symptoms of colorectal cancer, as described in NG12 or DG30. Studies were also included if the study recruited patients referred from primary care to secondary care. Studies reporting data for "Dual FIT" where patients are asked to provide two samples from different bowel movements were also included. All thresholds for defining a FIT were eligible for inclusion. The reference standard was not restricted but expected to comprise colonoscopy or computed tomography colonoscopy (CTC), other imaging tests or records follow up. Studies had to report diagnostic test accuracy data such that true positives, true negatives, false positives and false negatives could be extracted or calculated. Several additional outcomes were sought, including test failure rates, uptake of FIT, time to colonoscopy, time to diagnosis, and patient-reported outcomes such as health-related quality of life, preference and anxiety. Studies were also subgrouped according to several patient characteristics (anaemia, age, sex, ethnicity, medication that might affect FIT, blood disorders that might affect FIT).

The statistical synthesis pooled estimates of sensitivity and specificity at all reported diagnostic thresholds and provided summary estimates at all possible thresholds. The Evidence Assessment Group (EAG) did not use thresholds above or below those observed in the empirical studies. Studies were synthesised for each test separately. Sensitivity analyses investigated the effects of population type and reference standard, where data allowed. Where a statistical synthesis could not be performed, a narrative synthesis was conducted.

2.3.2 Cost-effectiveness methods

A mathematical model was constructed to simulate the experiences of patients presenting to primary care with symptoms of CRC. Three broad interventions were evaluated for individual tests which were: Intervention 1, the use of a single threshold for FIT for the full population to determine whether a person would be referred to the 2WW pathway or follow the safety netting pathway; Intervention 2, the use of two thresholds to determine if a patient would be referred to the 2WW pathway, referred to a intermediate pathway or would follow the safety netting pathway; and Intervention 3, which represented current practice with all patients at NG12 high/medium-risk being referred to the 2WW pathway with the remainder receiving a FIT and being subsequently assigned to the 2WW pathway or to the safety netting pathway. For the purpose of the economic model, the safety netting pathways was defined as consisting in one of four possible pathways being followed by patients: being referred to 2WW pathway anyway due to ongoing clinical concerns, being referred to non-urgent referral pathway (18WW), watchful waiting pathway, and being offered a second FIT (repeat FIT pathway).

The model was populated by published literature (synthesised by the EAG where appropriate), grey literature, estimates provided by clinical experts, and costs of FIT reported by the relevant companies. An initial decision tree model was used to categorise patients in terms of their true underlying disease status (and whether NG12 high/medium- or DG30 low-risk for intervention 3) and whether a FIT result was a true positive, a false positive, a true negative or a false negative. Following this, state-transition models were used to model patient survival, costs incurred and quality-adjusted life years (QALYs) gained. The model assumed that the proportional reduction in the total number of patients referred to the 2WW and 18WW pathway would translate directly into the equivalent reduction in time before diagnosis for patients in these pathways.

Outputs from this model included the life years gained (LYGs), QALYs and costs associated with each FIT strategy, the number of 2WW and 18WW referrals, the numbers receiving repeat FIT and allocated to the watch and wait pathway, the numbers of colonoscopies undertaken, and the mean time to a diagnosis of CRC, advanced adenomas (AAs,) and inflammatory bowel disease (IBD). To explore cost-effectiveness, incremental net monetary benefit was used as it allowed an easy comparison between FIT strategies which varied both by the specific FIT brand and the threshold(s) used to denote a positive, (intermediate in Intervention 2), or negative FIT result.

Eleven scenario analyses were performed, explored the impact of: i) decreasing the time of the underlying wait times associated with current care; ii) increasing the time of the underlying

wait times associated with current care; iii) assuming the loss of a full day's health for people receiving a colonoscopy; iv) assuming the loss of a full day's health for every month of delay associated with a definitive diagnosis for those in the 2WW or 18WW pathway and also for those with underlying bowel disease not in these pathways to account for patient anxiety whilst undiagnosed but with symptoms; v) assuming the use of dual FIT; vi) setting the prevalence of AAs and IBD to zero; vii) using a lower return rate for FIT; viii) assuming an alternative diagnostic accuracy of current FIT in low-risk patients in Intervention 3; ix) an increase in GP resource required for patients in the watch and wait and repeat FIT pathways without underlying disease from 1.9 appointments to 2.9 appointments; x) assuming an increased price associated with the FITs used in Intervention 3; and xi) assuming FIT to have perfect accuracy (sensitivity and specificity = 1.0) and return rate of 100%, to test an extreme scenario where no patients are missed by test or wrongly sent to 2WW.

2.4 Results

2.4.1 Clinical evidence results

49 studies were included in the review, across all tests and all subgroups and outcomes. Diagnostic test accuracy studies were categorised according to the recruitment criteria as either: population type 1 (studies closest to being a representative spectrum of all patients presenting to primary care with symptoms of CRC who meet NG12 or DG30 criteria); population type 2 (studies closest to being a representative spectrum of NG12 high/medium-risk patients); population type 3 (studies closest to being a representative spectrum of DG30 low-risk patients); or population type 4 (unclear/likely unrepresentative spectrum). The EAG had concerns about population type 4 studies, due to some studies including some patients on the basis of a positive FIT test in primary care, which may affect estimates of sensitivity and specificity.

There were risk of bias and/or applicability concerns with all the studies included in the review. Studies of type 2 or 4 generally scored high risk of bias for patient selection, since some primary care patients were not recruited, and studies of population type 1 or 3 generally scored high risk of bias for the reference standard, since not all patients received colonoscopy or CTC. Various other sources of bias were also noted.

The meta-analysis included data at all reported thresholds and provides summary estimates at all possible thresholds. There were only a small number of head-to-head comparative studies and so comparative test accuracy was not formally quantified. Considering a threshold of 10µg/g, the results were as follows, for sensitivity and specificity respectively: HM-JACKarc (n=16 studies), 89.5% (95% CrI: 84.6,93.4) and 82.8% (95% CrI: 75.2,89.6); OC-Sensor (n=11

studies), 89.8% (95% CrI: 85.9,93.3) and 77.6% (95% CrI: 64.3,88.6); FOB gold (n=3 studies), 91.2% (95% CrI: 68.2,99.8) and 80.3% (95% CrI: 64.9,91.1). No synthesis was conducted for QuikRead go, NS-Prime and IDK tests, since there was only one study for each. The estimates of sensitivity and specificity at 10µg/g respectively were: QuikRead go, 92.90% (95% CI: 68.5, 98.7) and 70.10% (95% CI: 66.1, 73.8); and NS-Prime, 71.40% (95% CI: 35.9, 91.8) and 83.60% (95% CI: 78.2, 87.9). The study of IDK Hb and IDK Hb/Hp only reported data at 2µg/g, and the sensitivity and specificity were calculated by IDK to be 87% (95% CI: 84.4, 89.6) and 88.1% (95% CI: 85.6, 90.6); IDK Hb/Hp, 82.6% (95% CI: 79.6, 85.6) and 80.8% (95% CI: 77.7, 83.9). As is usual for diagnostic test accuracy, sensitivity was higher at lower thresholds, and specificity lower at higher thresholds. No diagnostic test accuracy data was found for the combined use of OC-Sensor Ceres, IDK Hb + Hb/Hp or for IDK TurbiFIT tests. Equivalence data for IDK Hb and Hb/Hp, as well as for IDK Hb +Hb/Hp to the test used in the clinical study was not presented. Data on the equivalence of IDK TurbiFIT to IDK Hb and on the equivalence of OC-Sensor Ceres to OC-Sensor iO and PLEDIA was limited.

The sensitivity analyses showed that the exclusion of type 4 studies did not have a marked impact on the pooled estimates, with differences in the point estimates not consistent across the tests, and small in magnitude compared to the uncertainty (as quantified by the credible intervals and prediction intervals). In the analyses by population types 1, 2 and 3 separately, the summary estimates were similar and not statistically significant based on the overlap of the 95% credible intervals.

Four studies reported data using a Dual FIT strategy; two using HM-JACKarc, and one each using OC-Sensor and QuikRead go. In studies that reported estimates for both, the sensitivity was higher and specificity was lower when using Dual FIT (test interpreted as positive if either FIT positive) than that achieved when using only the first FIT test result to interpret the test.

Three studies compared two or more tests to each other in the same sample of patients (comparative diagnostic test accuracy studies). All three concluded there were some differences between tests, but none were able to conclude whether (and what) different FIT cut-off values would be required for each test.

Eleven studies reported diagnostic test accuracy for anaemic patients, three reported data according to age groups, three according to sex, and three for people taking medications which may affect FIT results. No studies were identified according to ethnicity or for people with blood disorders that may affect FIT results. Across these subgroup analyses, evidence was generally limited and sometimes inconsistent. It was not possible to conclude what or whether

different FIT thresholds may be required according to the patient characteristics specified in the NICE scope.

Eight studies reported data for the test accuracy of FIT tests for AA and IBD. Uncertainty was high in these analyses, with a large amount of heterogeneity between studies.

Ten studies reported test failure rates and these were largely between 2 and 5%. Only two studies reported test uptake in primary care and only one reported this where return of FIT was part of the diagnostic pathway. In this instance, the non-return rate was 9.4%. For Dual FIT, non-return rates appeared generally higher.

Two studies reported patient perspectives. The authors conclusions were that most patients found FIT acceptable, but strategies are needed to engage patients with more negative views of FIT, and shared decision making of patient and clinician should be considered for patients dissatisfied with relying on FIT results to decide on need for further investigation. Generalisability of these findings may have been affected by the fact that all patients included had been referred to secondary care.

One study reported on the impact of sociodemographic factors on FIT return rates and found higher return rates for females compared to males, older patients 65+ years compared to those <65 years, white patients compared to Asian, black and mixed/other ethnic groups, and the least socioeconomically deprived quintile compared to all other quintiles. Suggested strategies for addressing demographic differences in FIT return rate, which may reflect strategies for engagement with services as a whole, included following up after FIT non-return, using multiple languages, shared decision making and patient counselling to address concerns.

2.4.2 Cost-effectiveness results

For the vast majority of FIT strategies, the incremental net monetary benefit iNMB was positive compared with current care regardless of the cost-effectiveness threshold used, or whether one, or two, thresholds were used. These conclusions were robust to the sensitivity analyses undertaken. The iNMB were typically in the range of £200 to £350 per patient which were driven by the reduction in the costs of colonoscopy, although there was a slight decrease in patient health predominantly due to patients who had a false negative FIT and who would have received a colonoscopy under current practice. Given the uncertainty in model parameters and in the inherent simplification associated with reducing a complex real-world problem into a mathematical model the EAG did not feel comfortable providing a robust estimation of which FIT brand and which threshold(s) the iNMB was highest. It is anticipated that a more nuanced

discussion would be undertaken at the NICE Appraisal Committee with the input of specialist committee members than could be incorporated robustly within the modelling undertaken.

2.5 Discussion

The systematic review identified diagnostic test accuracy data for 7 of the 9 tests. Only one study each was identified for QuikRead go, NS-Prime, IDK Hb and IDK Hb/Hp and in all these cases the studies were relatively small (n analysed <700, CRC events <25), and subject to limitations. No diagnostic test accuracy data was identified for OC-Sensor Ceres, IDK TurbiFIT and IDK Hb+Hb/Hp. Data on the equivalence of the IDK tests to the test used in the clinical trial was not presented. Data on the equivalence of OC-Sensor Ceres to OC-Sensor iO and PLEDIA were limited. The statistical synthesis produced summary estimates of sensitivity and specificity across all possible thresholds where data allowed. There was insufficient data to conduct an analysis of the comparative diagnostic test accuracy between tests. Dual FIT studies were only identified for three tests (HM-JACKarc, OC-Sensor, QuikRead go). There was insufficient and sometime contradictory data relating to patient characteristics (anaemia, age, sex, ethnicity, medication that might affect FIT, blood disorders that might affect FIT), and no conclusions could be drawn on whether different thresholds should be used in people with these characteristics. FIT was found to be generally acceptable but return rates may be different according to sociodemographic factors. There were limitations to both the evidence base and the systematic review that should be taken into consideration when interpreting the evidence.

For all FIT brands there are strategies at which the iNMB is positive compared to current care. The exact brand and threshold(s) which generate the greatest iNMB (at a selected threshold) could not be robustly determined due to the similarity of iNMB values, parameter uncertainty and the possibility of omissions from the model structure.

2.6 Implications for service provision

The model makes assumptions about the effects of safety netting, which may not be consistent with the safety netting offered across the country at present. Standardisation of and improvement to safety netting practice may be required. Interventions may be required to increase FIT return, especially in some socioeconomic groups, and to improve the experience of a minority who have negative views about FIT, and dissatisfaction with reliance on FIT for diagnostic purposes.

2.7 Suggested research priorities

Research priorities include investigating the comparative diagnostic test accuracy between tests, and whether different thresholds are required for patients with characteristics that may affect FIT accuracy. It is likely that new primary studies would be required to enrich the

evidence base. Whilst the analysis was not able to detect an effect of population type, enrichment with FIT positives or the reference standard used, these are all issues that should be considered in future primary studies and evidence syntheses since the analyses conducted here were not conclusive.

3 BACKGROUND AND DEFINITION OF THE DECISION PROBLEM

3.1 Condition and aetiology

3.1.1 Aetiology, pathology and prognosis

The aetiology of colorectal cancer (CRC), like many cancers, is multifactorial and involves an interplay of hereditary, environmental and lifestyle factors. Around two-thirds of cases occur in those with no hereditary predisposition, and are caused by a wide range of genetic mutations and epigenetic aberrations that may occur as a result of potentially modifiable risk factors.[1] A history of CRC in the family is evident in around 25% of cases, with around 5% attributable to hereditary cancer syndromes.[1 2]

Colonic polyps are abnormal growth in the lining of the bowel. These are usually asymptomatic and are a common finding during colonoscopy. Although most polyps do not become cancerous, most CRCs arise form colonic polyps and their removal significantly reduces the risk of CRC. There is a greater risk of progression to CRC in people with large and/or multiple polyps, but this usually takes many years. [3] Therefore, an incidental but important consequence of colonoscopic investigations for CRC may be the opportunity to identify and remove polyps.

The prognosis of CRC is dependent on disease stage. Most people with early CRC can be cured but late-stage disease is associated with a low 5-year survival rate. Therefore, early identification is desirable.

3.1.2 Epidemiology and incidence

CRC is the fourth most common form of cancer in the UK. Approximately 42,000 new cases of CRC are diagnosed each year, resulting in around 16,800 CRC-related deaths annually.[4] The Global Burden of Disease study[5] estimates that there were 1.8 million (95% confidence interval (CI) 1.8-1.9) incident cases globally of CRC in 2017, with an age-standardised rate of 23.2 (95% CI 22.7 to 23.7) per 100,000 person-years, an increase of 9.5% since 1990. The regions with the highest incidence were Australasia, high-income Asia Pacific, high-income North America and Europe. Incidence was higher in men than women in all regions.

CRC is predominantly a disease of older adults, though in recent years incidence has increased sharply in younger adults aged 20-39.[6] Historically a disease of affluence in the UK, the influence of socioeconomic factors has also changed in recent years, as surveillance data showed an increased risk for adults from areas with higher deprivation between 1996-2010 for men,[7, 8] and in the 2010s for women.[9]

3.1.3 Burden of disease

The Global Burden of Disease study estimates that 0.33 disability-adjusted life years (DALYs) were lost in the UK per 100,000 person-years in 2019, a number that has fallen since the 1990s when it was estimated to be 0.48 per 100,000 person-years.[5]

3.2 Current service provision

3.2.1 National guideline 12 (NG12) high/medium-risk and diagnostic guideline 30 (DG30) low-risk patients

National guideline 12 (NG12) describes the diagnostic pathway for patients presenting to primary care with symptoms suggestive of CRC[10] (Figure 1). Within this guideline, patients with the following symptoms (referred to as NG12 high/medium-risk patients in this assessment) should be referred to secondary care with an urgent 2 week wait (2WW) suspected CRC referral. NG12[10] states:

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

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

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

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

In July 2017, NG12[10] was partially updated by Diagnostics guidance 30 (DG30).[11] In this update, the guaiac faecal occult blood test (fOBT), which had been recommended for use in low-risk patients, was replaced with FIT; hence, DG30[11] recommends use of faecal

immunochemical tests (FIT) for suspected CRC in people without rectal bleeding who have unexplained symptoms but do not meet the criteria for a suspected cancer pathway referral - that is - they are at low-risk of CRC (referred to as DG30 low-risk patients in this assessment). These patients include:

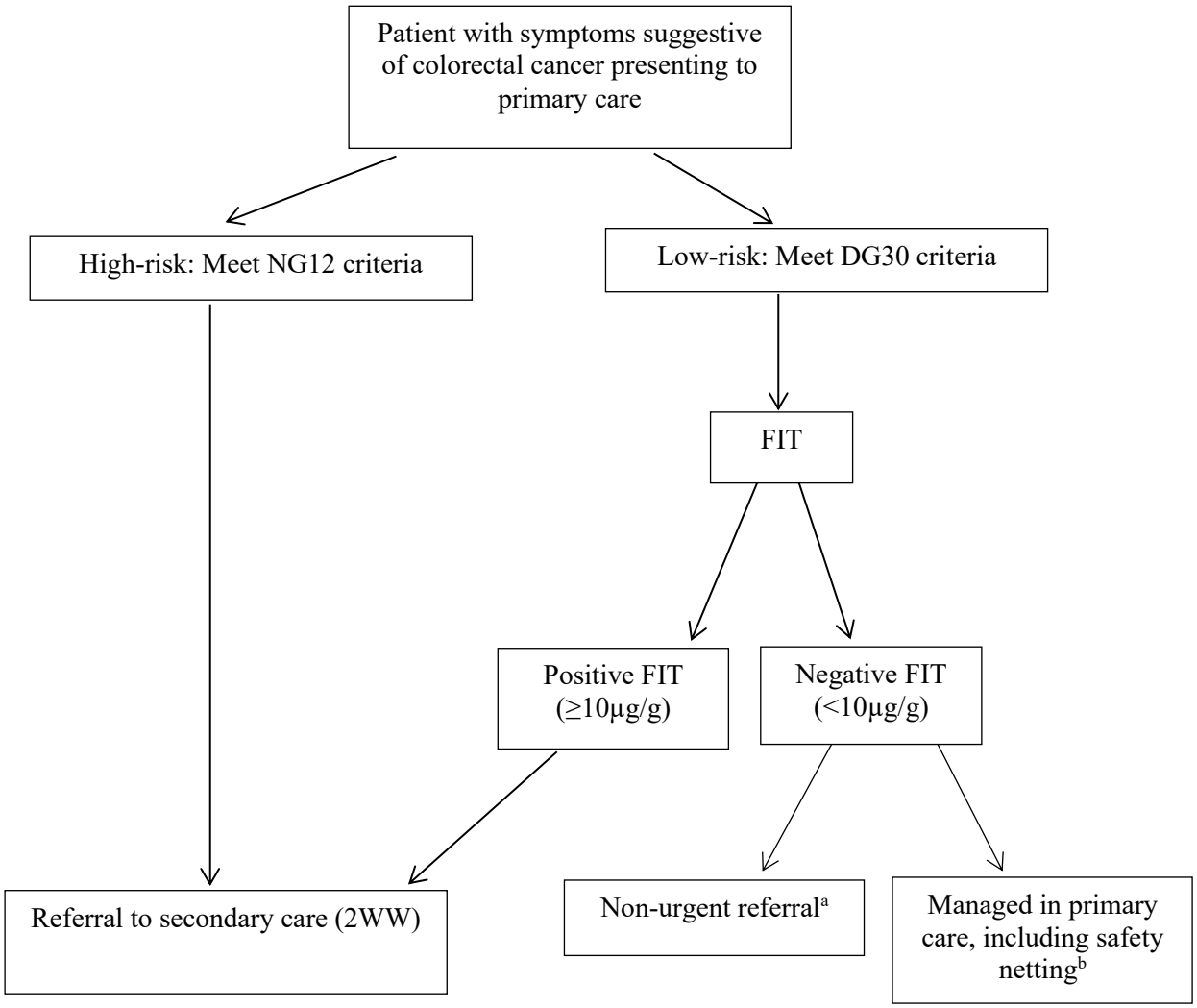
- People aged 50 and over with unexplained:
    - abdominal pain or
    - weight loss, or
- People aged under 60 with
    - changes in their bowel habit or
    - iron-deficiency anaemia or
- People aged 60 and over and have anaemia even in the absence of iron deficiency.

The review undertaken for DG30 showed that the specificity of FIT was high (>90%), hence it was recommended for use as a rule in test at a threshold of 10 micrograms haemoglobin per gram of faeces (µg Hb/g, hereafter referred to as µg/g). Patients testing positive by FIT should be referred on to the 2WW suspected CRC pathway.

What happens in secondary care following referral is thought to vary across England: it may be to a specialist who will order further tests (colonoscopy, CTC, or other tests as they see fit) or may be a direct referral by a GP to an imaging test such as colonoscopy or CTC. The choice of imaging test may be dependent on local practice guidelines or age and comorbidities that contraindicate colonoscopy. CTC may be necessary where colonoscopy fails or is inappropriate. Colon capsule endoscopy is a relatively new imaging modality, whereby a small capsule containing a camera is swallowed in order to image the digestive tract, and is used in some areas of the UK. During colonoscopy, a biopsy may be taken for histological confirmation, unless this is contraindicated (e.g., blood clotting disorders).

It is recommended that patients testing negative by FIT are followed up in primary care. This should include “safety netting” as described for all cancer pathways in NG12, to avoid missing disease (cancer or otherwise) in people with negative FIT results (see Section 3.3.8.5).[10] Safety netting in NG12 includes an awareness of the possibility of false-negatives, and re-testing either after a period of time or upon the emergence of new symptoms, or the recurrence, persistence or worsening of existing symptoms.[10] Safety netting may also include strategies for diagnosing other gastrointestinal conditions such as inflammatory bowel disease (IBD, a term used to describe Crohn's disease (CD) and ulcerative colitis (UC)), and further monitoring for colorectal or other types of cancer.

Figure 1: The diagnostic pathway for patients presenting to primary care with symptoms of colorectal cancer. Based on NG12[10] and DG30[11]



2WW, two week wait; DG30, diagnostic guideline 30; FIT, faecal immunochemical test; NG12, National Guideline 12; t, threshold

[a] *non-urgent referral is part of current care for some FIT negative patients, based on clinical judgement*

[b] *Safety netting is discussed in Section 3.3.8.5*

3.2.2 Speciality guidance during the height of the coronavirus pandemic

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

urgent colonoscopy or computerised tomography (CT) which could be CTC or plain CT. People with between 10 and 100 µg/g, or people with more than 100 µg/g who have had a colonoscopy requiring no further investigation in the last 3 years, would be referred for prioritised colonoscopy or colonic imaging (CTC, plain CT, or colon capsule endoscopy). People with less than 10 µg/g would be managed using safety netting processes (see Section 3.3.8.5).

3.2.3 ACPGBI/BSG guideline and NHS England letter

In 2022, the Association of Coloproctology of Great Britain and Ireland (ACPGBI) and the British Society of Gastroenterology (BSG) published guidance on FIT in patients with signs or symptoms of suspected CRC (ACPGBI/BSG guidance).[12] This guidance was based on a systematic review of the available evidence, expert opinion and was agreed by consensus. Economic evaluation was not conducted. In October 2022, NHS England published letters[13, 14] endorsing the use of the ACPGBI and BSG guidance on FIT in primary care, stating it should be implemented in full.

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

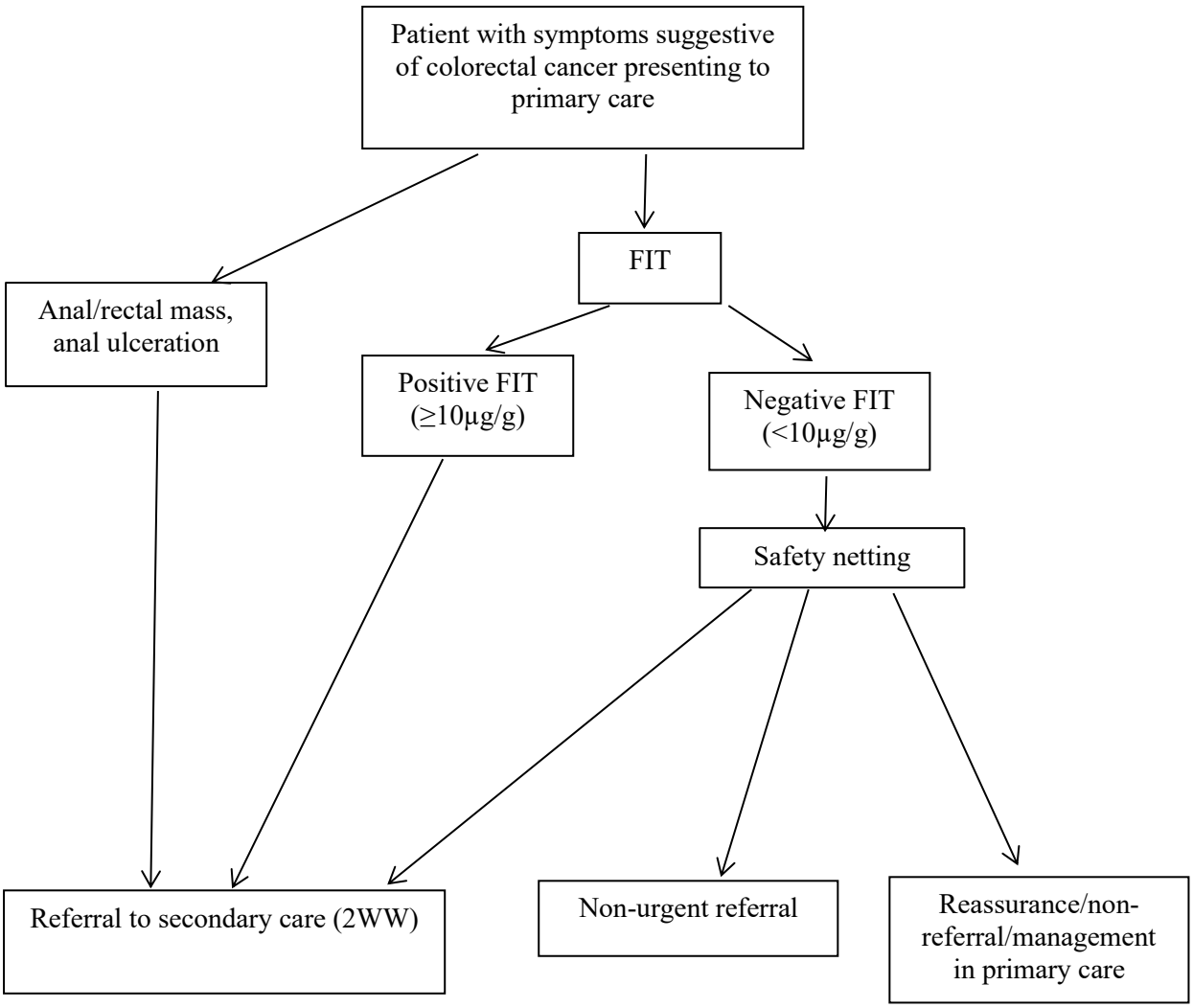
Figure ***2***.

The ACPGBI/BSG guideline notes that FIT should not be the sole determinant of referral. Patients without symptoms were not considered in the guideline and should not be referred on the basis of a positive FIT, except within the context of the national screening programme. Patients with negative FIT should not be excluded from referral; where FIT is <10µg Hb/g but there are persistent and unexplained symptoms which concern the GP, the patient should be referred to secondary care for evaluation. This referral may be to routine or urgent pathways, but not necessarily to the CRC pathway. Those with abdominal mass should be referred and a FIT ordered at the same time for use in secondary care. Those with anal/rectal mass or anal ulceration should be referred on the urgent 2WW suspected CRC pathway without a FIT.

The NHS England letter also contains recommendations on safety netting for people with negative FIT results. This is discussed in more detail in Section 3.3.8.5.

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

Figure 2 The diagnostic pathway for patients presenting to primary care with symptoms of colorectal cancer as recommended in the ACPGBI/BSG guideline[12]



2WW, two week wait; FIT, faecal immunochemical test; µg/g, micrograms of faecal haemoglobin per gram of faeces

3.3 Description of the decision problem

3.3.1 Purpose of the decision to be made

Early diagnosis and treatment of CRC in people presenting to primary care with symptoms can improve survival and cure rates. The NICE NG12 Guidelines[10] introduced in 2015 expanded the symptoms-based referral criteria recommended in NICE Clinical Guideline 27 (CG27) (2005, now unavailable) to a wider set of symptoms. This resulted in an increase in the number of 2WW referrals, but there was not a corresponding increase in the proportion of patients investigated who have cancer.[15] Indeed in 2018, of 392,588 referrals made with suspected cancer on the 2WW pathway in England only 13,168 (3.3%) had a cancer. In addition, in August 2022, 28% of people seen by a specialist for suspected CRC were not seen within 2 weeks of urgent referral, and 53% did not have a diagnosis within 28 days (NHS cancer waiting times, August 2022). Of 15,053 people treated for lower gastrointestinal cancer in 2020-21 under a suspected cancer pathway referral, only 50.6% received treatment within 62 days following an urgent GP referral (compared with an operational standard of 85%).

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

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

NICE heard via consultation with stakeholders and the NHS that the current symptom-based referral pathway, using the NG12 and DG30 criteria, is difficult for GPs to implement. The ACPGBI/BSG guideline[12] and the meta-analysis that informed the guidelines[16] also found that there is no clinically significant difference in sensitivity when FIT is used in patients presenting with DG30 and NG12 symptoms as well as those presenting with certain individual symptoms (rectal bleeding, iron deficiency anaemia and abdominal pain), though this guideline did not consider the impact on cost-effectiveness.

There is evidence that FIT is a better predictor of CRC risk in patients than symptoms alone[17] and could result in fewer referrals of people without CRC to the urgent 2WW suspected CRC

pathway. Therefore, triage with FIT could mean that people who are unlikely to have CRC may avoid unnecessary referral for investigations. They would also avoid the associated disadvantages of referral such as patient anxiety, time off work and loss of economic productivity, as well as the rare adverse events associated with colonoscopy for example, bleeding, perforation and death. Furthermore, those that are likely to have CRC can be prioritised more effectively[18] leading to a reduction in time to diagnosis. This may also release colonoscopy capacity to allow people on non-urgent referral pathways to be seen more quickly. The extent to which colonoscopy capacity is released will be dependent in part on the threshold used to define a positive test for the symptomatic patients.

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

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

3.3.2 Place of the intervention in the treatment pathway

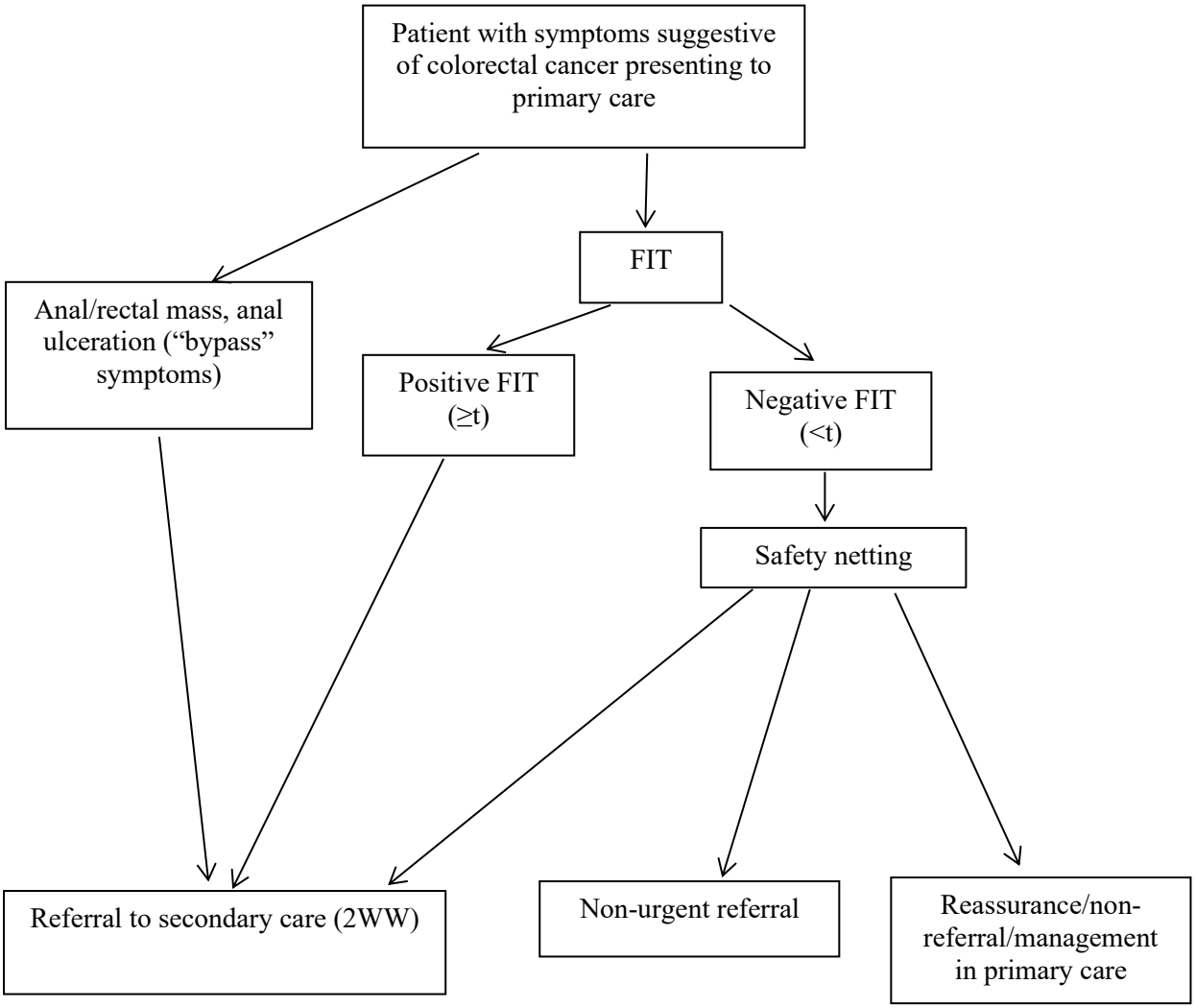
This assessment considered the use of FIT in people presenting to primary care with gastrointestinal symptoms indicating a risk of CRC. The treatment pathway and proposed position for FIT are shown in

Figure **3**.

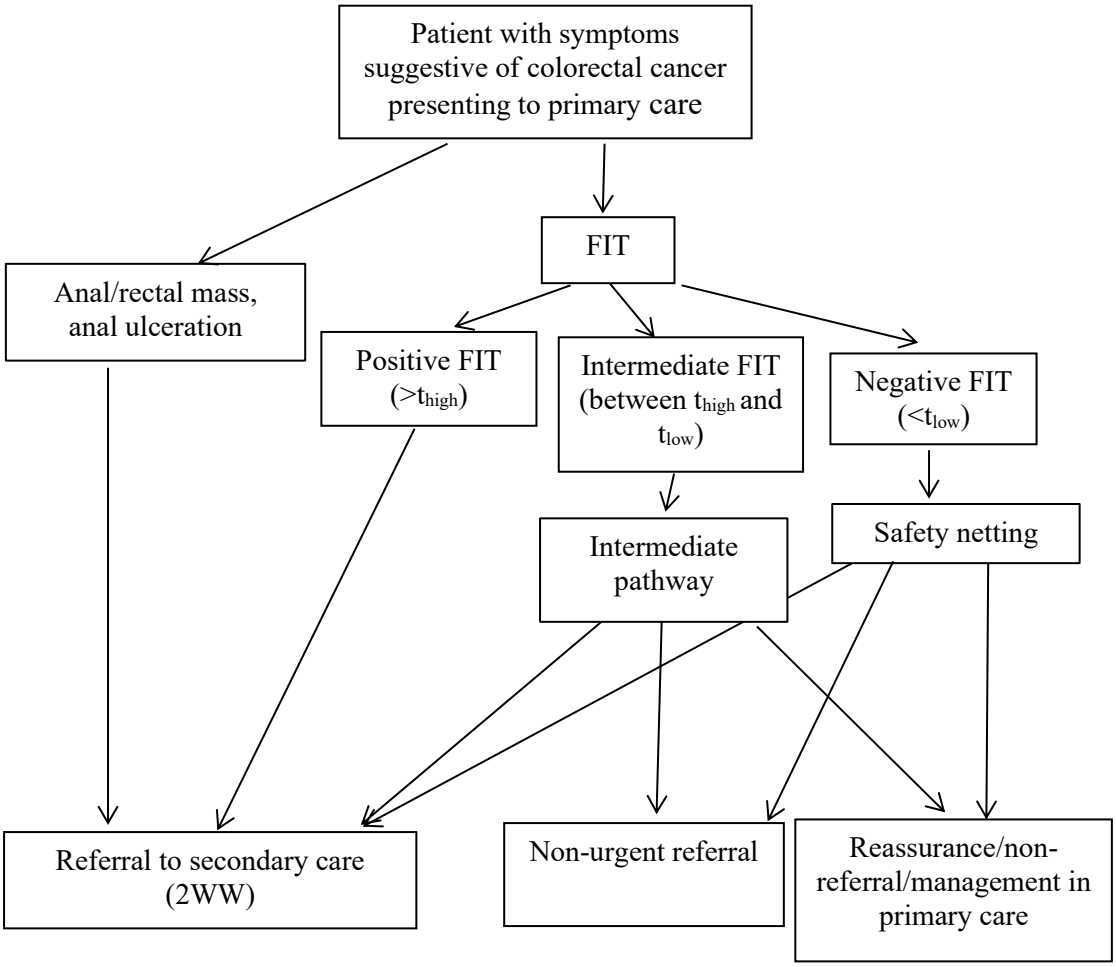
FIT was to be evaluated as an adjunct to clinical assessment to guide referral of a symptomatic population to the suspected CRC pathway. Consistent with the ACPGBI/BSG guideline, this population included both those meeting NG12 criteria for an urgent 2WW suspected CRC referral, and those meeting DG30 criteria for a FIT test, and excluded those with rectal or anal mass, or anal ulceration (who should go straight to urgent 2WW suspected CRC referral, termed "bypass symptoms" in this assessment). Patients would receive the test in primary care, and the result of the test would be used to determine who would proceed to secondary care and who would be followed up in primary care with safety netting.

Figure 3: Proposed new pathway incorporating FIT for all patients in primary care: A) using a single FIT threshold; and B) using two FIT thresholds to create an intermediate risk group who would follow a different diagnostic pathway

A)



B)



2WW, two week wait, FIT, faecal immunochemical test; t, threshold; t_{high}, higher threshold; t_{low}, lower threshold

3.3.3 Definition of the intervention

Quantitative faecal immunochemical tests (FIT) are designed to detect occult (small amounts) of blood in stool samples (faecal haemoglobin) using antibodies specific to human haemoglobin.

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

Immunoturbidimetric FIT contains particles which are coated in antibodies specific to human haemoglobin. Six of the tests within the scope of this assessment use this methodology (see Table 1). The antibodies bind to haemoglobin present in the faecal sample creating complexes

which are detected using turbidimetry (how much light is absorbed when passed through a solution).

ELISA FIT uses antibodies specific to human haemoglobin to bind haemoglobin in the faecal sample to the surface of microtiter wells. Two of the tests within the scope for this assessment use this methodology (see Table 1). This is then treated with chemicals to produce a colour change. The intensity of the colour is proportional to the amount of haemoglobin in the sample. Some assays may also include antibodies for human haptoglobin. Haptoglobin is a protein produced by the liver which binds to haemoglobin making it less likely to break down during transit through the gastrointestinal tract. The detection of haptoglobin is claimed to increase the likelihood of detecting lesions higher in the colon.

Different FIT tests may report outcomes using either the concentration of haemoglobin in the sampling device buffer (nanograms Hb/mL buffer) or as concentration of haemoglobin by mass of faeces (μg/g). As the amount and type of buffer used varies between manufacturers, the World Endoscopy Organization's expert working group on FIT for CRC screening recommended that μg/g should be used as a standard measure that can be compared easily between tests.[19]

3.3.3.1 Strategies and thresholds for using FIT as a triage tool

Since the test is quantitative, thresholds may be varied to achieve optimal clinical and cost effectiveness outcomes with respect to colonoscopy capacity, quality-adjusted life years (QALYs) or net monetary benefit (NMB).

Strategies using one FIT threshold were to be investigated, where FIT above a threshold resulted in referral to the urgent 2WW suspected CRC pathway, whilst FIT below the threshold would result in safety netting (see Sections 3.3.8.5 and 5.3.3.1). A strategy using two FIT thresholds (t_1 and t_2) was also to be considered (see B) and is described in Sections 3.3.8.1.1 and 5.3.3.1. A strategy using two FIT tests (Dual FIT) was also of interest (see Sections 3.3.8.4 and 5.3.5.1).

There are several FIT tests within the scope of this assessment. These are described in Sections 3.3.3.2 to 3.3.3.8 and are summarised in Table 1.

3.3.3.2 HM-JACKarc system

The HM-JACKarc system (Minaris Medical Co., Ltd) is a fully automated quantitative immunoturbidimetric FIT system. The system comprises a sample collection device (designed

to measure 2 mg of faeces) which contains 2 mL of stabilising buffer, latex agglutination reagent, and buffer solution. The assay is compatible with the HM JACKarc analyser, which can process up to 200 samples per hour, with a maximum capacity of 80 samples per run.

3.3.3.3 FOB Gold

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

3.3.3.4 OC-Sensor

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

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

MAST Diagnostics state that the OC-Sensor iO will soon be replaced by the OC-Sensor Ceres.

3.3.3.5 NS-Prime

The NS-Prime (Alfresa/Abbott) is an automated quantitative immunoturbidimetric FIT system. The NS-Prime comprises a specimen collection container which collects 10 mg of faeces in 1.9 mL of buffer solution (Carroll *et al*. 2014).[20] The test is run on the NS-Prime clinical chemistry analyser.

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

3.3.3.6 IDK TurbiFIT

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

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

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

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

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

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

3.3.3.8 QuikRead go iFOBT

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

Table 1: Summary of interventions (adapted from Table 1 from the NICE scope[21])

Test	**Test principle**	**Analyser compatibility**	**Sample size required** (mg)	**Measuring range** (µg/g)	**Limit of detection** (µg/g)	**Limit of quantitation** (µg/g)	**Throughput**
HM-JACKarc	Immunoturbidimetry	HM JACKarc analyser	2	7 to 400	2	7	200 samples per hour
FOB Gold	Immunoturbidimetry	Various	10	Varies according to the analyser used	Varies according to the analyser used	Varies according to the analyser used	Dependent on the analyser used
OC-Sensor	Immunoturbidimetry	OC-Sensor PLEDIA	10	2 to 50,000	2	2	320 samples per hour
	Immunoturbidimetry	OC-Sensor iO	10	2 to 200	2	4	88 samples per hour
NS Prime	Immunoturbidimetry	NS-Prime analyser	10	4 to 240	4	10	300 tests per hour
IDK TurbiFIT	Immunoturbidimetry	Various	15	Varies according to the analyser used	Varies according to the analyser used	Varies according to the analyser used	Dependent on the analyser used
IDK Hemoglobin ELISA	ELISA	Various (ELISA plate reader with a photometer (Dynex DS2 and DSX systems))	15	0.18 to 50	0.15	0.18	Dependent on the analyser used
IDK Hb/Hp complex ELISA	ELISA		15	0.25 to 50 µg Hb/Hp/g	0.16 µg Hb/Hp/g	0.25 µg Hb/Hp/g	Dependent on the analyser used
QuikRead go iFOBT (point-of-care test)	Immunoturbidimetry	QuikRead Go analyser	10	10 to 200	2.5	9.5	Less than 2 minutes per test.

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

3.3.4 Populations and relevant subgroups

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

Certain symptoms may indicate patients should be referred directly to the urgent 2WW suspected CRC pathway (people with palpable rectal or anal mass or anal ulceration, termed "bypass symptoms" in this assessment) and these patients were excluded from the scope. In contrast to DG30 (see Section 3.3.5), rectal bleeding was not considered a symptom that would preclude the use of FIT, since clinicians indicated during the scoping process for this assessment that FIT tests can be used in those with rectal bleeding.

There are reports that suggest that faecal haemoglobin levels may differ according to certain patient characteristics. If confirmed, different cut-off values may be needed according to the following characteristics:

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

This assessment aimed to identify diagnostic test accuracy data within these subgroups to help inform whether alternative thresholds may be required to achieve accuracy equivalent to that for patients without these characteristics. Economic modelling was not planned for these subgroups.

3.3.5 Relevant comparators

Current practice corresponds to standard care according to current NICE guidelines NG12 and DG30 (see Section 3.2.1). This includes:

- Clinical assessment and referral for further investigation in secondary care
- Use of FIT (threshold of 10 μg/g) to guide referral only for those with 'low-risk' symptoms without rectal bleeding (in line with NICE DG30).

Feedback from clinical experts and stakeholders during the scoping stage of this assessment was that stratification by symptoms is a poor predictor of risk of CRC. Any resulting guidance that differentiates

between the risk groups currently defined in NICE guidance would not address this problem. Therefore, despite the possibility of differential cost-effectiveness by subgroup, NICE's scope[22] stated that the intervention arm should not subgroup according to NG12 high-risk and DG30 low-risk categories and should not exclude those with active rectal bleeding, to avoid making recommendations according to symptom-based criteria. Consequently, the comparator was a blended group of people who would currently be considered under the guidance of NG12 and DG30.

The NICE scope noted that the comparators for the modelling may differ.

3.3.6 Healthcare setting

The assessment related to the use of FIT in primary care.

3.3.7 Outcomes

The NICE scope[22] states that intermediate outcomes of interest may include:

- Diagnostic accuracy at different FIT thresholds for CRC, AA and IBD
- Risk of CRC (and IBD and AAs) in relevant subgroups according to FIT threshold
- Test failure rates
- Prognostic implications of false-negative results
- Uptake (completion) of FIT in primary care
- Number/proportion of people referred to secondary care
- Number/proportion of people followed up in primary care
- Duration of validity of negative test (implications for follow-up)
- Number/proportion of urgent (2WW suspected cancer) specialist appointments
- Number/proportion of urgent (2WW suspected cancer) colonoscopy/CTCs
- Number/proportion of non-urgent colonoscopy/CTCs
- Time to colonoscopy/CTC
- Time to diagnosis of CRC or other conditions
- Number/proportion of colonoscopy/CTCs that do not detect CRC
- Number/proportion of colonoscopy/CTCs that do not detect significant bowel pathology
- Number/proportion of people presenting to emergency departments with symptoms of CRC.

The NICE scope[22] states that clinical outcomes for consideration may include:

- Number of CRC diagnoses
- Number/proportion of CRC diagnoses from urgent referrals
- Stage of detected cancers
- Number/proportion of people identified with other bowel pathologies
- Number/proportion of people with AAs detected, or detected and treated

- Morbidity including adverse events associated with colonoscopy
- Mortality.

The NICE scope[22] states that patient-reported outcomes for consideration may include:

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

The NICE scope[22] states that costs were to be considered from an NHS and Personal Social Services (PSS) perspective. Costs for consideration included:

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

A lifetime horizon was to be used. The cost-effectiveness of FIT versus usual practice was to be expressed in terms of the incremental cost per QALY gained (ICER). Net health benefit (NHB) was to be used when comparing multiple interventions, but the EAG has presented NMB to aid the committee's interpretations of the results of the economic analyses.

3.3.8 Other considerations

There is known to be heterogeneity within care pathways across the country and this was to be investigated in the assessment.

3.3.8.1 FIT threshold for referral

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

- To optimise the treatment pathway for clinical effectiveness (QALYs) or cost-effectiveness (in terms of NHB) and to investigate impact on numbers/proportions of referrals
- Because faecal haemoglobin levels are thought to differ according to certain patient characteristics (see Section 3.3.4), different cut-off values may be needed for these subgroups to avoid potential equity issues

Both reasons for threshold alteration were to be considered in the assessment.

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

Two FIT thresholds could be used to define low (FIT lower than $t_{low)}$, intermediate (FIT between t_{low} and t_{high}), and high-risk populations (FIT >t_{high}). In this strategy, people in the intermediate risk group (with FIT between t_{low} and t_{high}) may have more intensive monitoring of their condition than in the low-risk group or be referred to a specialist safety netting pathway (see Section 3.3.8.5). The management pathway for the intermediate group was unclear and was addressed in the modelling.

3.3.8.2 Measurements and diagnostic test accuracy of different tests and analysers

Different tests, different analysers and different combinations of tests and analysers (see Table 1) may have different measuring ranges, may give different absolute measurements, and may have different test accuracy. The NICE scope[22] notes that accuracy should be analysed according to the test-analyser combination. This was considered in the clinical review of evidence.

3.3.8.3 Use of FIT alongside bypass referral

As already noted, clinical experts advised NICE that rectal bleeding would no longer be considered a reason to bypass FIT. Both the NHS England letter and the ACPGBI/BSG guideline stated that presence of a palpable rectal or anal mass, or anal ulceration were symptoms that indicated that patients should move straight to a 2WW referral, thereby bypassing FIT. Some clinical experts said that FIT could still be useful alongside referral, to help choose the method of further investigation, and may be required by some secondary care centres. Since the bypass symptoms are not part of the decision problem population, this assessment did not include them in the modelling.

3.3.8.4 Dual testing

Two FIT tests can be used to guide referral. There are two main ways in which two tests can be used, and in this assessment, these are termed “Dual FIT” and “Repeat FIT”.

“Dual FIT” was defined in the NICE scope[22] as using two samples from different bowel movements rather than a single sample from one bowel movement. The scope notes that it is different to using FIT as part of a safety netting programme, which we are calling in this assessment “Repeat FIT”. Repeat FIT has also been defined elsewhere as referring to the use of a second FIT after a decision to refer or not refer has been made.[23] Repeat FIT usually takes place weeks or months later (see Section 3.3.8.5) as a result of continuing or worsening symptoms, whereas Dual FIT is given to all patients on the basis of their initial consultation.

This assessment considered Dual FIT as a testing strategy. Based on clinical expert opinion, people would be referred to the suspected cancer pathway if either FIT sample were positive. Dual FIT may result in fewer false-negative results, more false-positive results and higher costs of FIT testing.

Repeat FIT is considered as part of safety netting within the economic modelling for this assessment (see Section5.3.4.4). Studies reporting data on Repeat FIT are reported in Section 4.3.14.1.

3.3.8.5 Safety netting

Clinically, safety netting refers to various strategies and processes used in the diagnostic pathway to avoid missing disease (cancer or otherwise). In the context of the CRC pathway, this is most usually for those who do not get initially referred to secondary care. This section outlines some of the available recommendations on safety netting, which clinical advisors to the EAG indicated are implemented to differing extents across the country.

DG30 modelling assumed the following for safety netting (persons with 'negative FIT'): (i) if they had cancer, they would have a delay in diagnosis of less than 12 months as they would re-present with continuing or worsening symptoms, (ii) for those without cancer a proportion would also have persistent symptoms, some of whom would receive colonoscopy, and some would receive a repeat FIT. For (ii), proportions were estimated based on clinical opinion (two clinicians who provided quite different estimates, Table 26 of DG30[24]); the DG30 EAG assumed 32.5% of patients who tested negative with FIT/gFOBT, would persist in their symptoms and would to receive colonoscopy and 20% had repeat FIT.

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

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

- "*Providing the patient with clear information about who to contact if they develop new symptoms or if their existing symptoms worsen.*
- *Using advice and guidance via eRS* [electronic referral system] *to guide management of patients with persistent or troublesome symptoms.*
- *Offering a second FIT if ongoing clinical concerns remain.* [NB, this is called "Repeat FIT" in this assessment]

- *Referral to a non-specific-symptoms urgent cancer pathway, if appropriate and there are ongoing concerns about possible cancer.*
- *Management of FIT negative patients in an outpatient setting following referral on a non-urgent pathway. For example, the North Central London Cancer Alliance has developed a FIT negative, non–urgent referral pathway, as has Oxford University Hospitals NHS Foundation Trust.*"

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

Safety netting was to be included as part of the diagnostic pathway of patients with negative FIT results in this assessment, exploring different assumptions about its composition (see Sections 5.3.3.1 and 5.3.4.4).

3.3.8.6 Other conditions with gastrointestinal symptoms

Patients presenting with symptoms of CRC may have other gastrointestinal pathologies such as IBD (CD or UC), diverticular disease or AAs. Colonoscopy is required to diagnose IBD, and to identify and treat AAs.

The COLOFIT project conceptual modelling has opted to explicitly include IBD (both CD and UC) within the model due to the known impact of a delayed diagnosis on prognosis, costs and quality of life. Other bowel diseases were not modelled explicitly due to a lack of clarity around whether a diagnostic delay is likely to cause harm. In this assessment, a similar approach was taken, and IBD and AAs were included within the scope of the modelling; the pathways for these patients were also considered to have an impact in outcomes due to a delay in diagnosis. A delay in diagnosis for IBD may worsen quality of life and patient outcomes, whilst AAs are largely asymptomatic and colonoscopic findings in AAs are largely incidental, but some may eventually develop into CRC if not treated, which may have an impact on patients' lifetime survival, health-related quality of life (HRQoL) and costs (see Section 5.3.3).

3.3.8.7 Urgent 2WW suspected CRC pathway and secondary care management

Clinical advisors indicated during the scoping process for this assessment that there was heterogeneity in current practice regarding what happens in secondary care upon a referral to the urgent 2WW suspected CRC pathway. This was to be appropriately represented within the project.

3.3.8.8 Non-urgent referral pathway

Clinical advisors indicated during the scoping process for this assessment that there was heterogeneity in current practice regarding what the non-urgent referral pathway entails. This was to be appropriately captured within the project.

3.3.8.9 Non-completers of FIT tests

A proportion of patients do not return their FIT tests. Based on the systematic review conducted for DG30, FIT was returned by 41% (in a study where patients were sent an invitation to participate along with their referral letter) to 98% (in a study where patients were given the specimen collection device at their initial consultation with a gastroenterologist) for patients using OC Sensor, and 56%-66% patients using HM-JACKarc. This was to be taken into account within the project.

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

Evidence on equivalence of tests and test-analyser combinations (e.g., Bland Altman plots) was not sought or statistically synthesised by the EAG. Evidence submitted by companies relating to equivalence was to be considered by the EAG to inform modelling scenarios.

Development of a risk prediction model using FIT and clinical characteristics was not within the scope of the assessment. This type of work was being conducted by other groups (e.g., the NICE FIT group, COLOFIT). A review of risk prediction models is also not within the scope of this assessment, since this work is being conducted by the COLOFIT group.

4 CLINICAL EVIDENCE

4.1 Methods

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

Summary of the approach to the review: The ACPGBI/BSG guidance[25] was based on a recent systematic review of the literature relating to clinical efficacy and diagnostic test accuracy.[26] Some of the authors of that review were clinical advisors to the EAG (Mr Muti Abulafi; Mr Kevin Monahan; Dr Richard Booth; Dr Rachel Carten) and they shared their review work as a basis for the review for this assessment. There were some notable differences in scope between the review for this assessment and the ACPGBI/BSG, namely that a limited number of thresholds were eligible for inclusion in ACPGBI/BSG, different subgroup analyses were planned, and the focus was not on recruitment of patients in primary care only. To ensure that all threshold and relevant subgroup data was identified, the list of ACPGBI/BSG excluded studies was scrutinised to identify studies relevant to this assessment, and where data were not extracted for all thresholds and subgroups reported in a study, the original study was revisited to perform *de novo* data extraction as detailed in Section 4.1.8. Studies not relevant to this assessment that were included in the ACPGBI/BSG review were excluded.

The protocol for this review was registered on the International prospective register of systematic reviews (PROSPERO, registration number CRD42022383580).

4.1.1 Population

Studies were included if they recruited people presenting to primary care with signs or symptoms indicating a risk of CRC. Signs and symptoms of CRC were defined as those described in NG12 and DG30 (see Section 3.2.1), though studies were not excluded if recruitment criteria were wider than those listed in NG12 and DG30, or were narrower. Studies reporting data relating to the subgroups specified in the population section (Section 3.3.4) of the decision problem (e.g., age, sex) were also included (hereafter these are called "patient characteristics studies"), but studies reporting on very narrow populations that did not relate to a subgroup of interest, such as those with rectal bleeding only, were excluded. Studies which did not recruit only patients presenting to or referred from primary care (e.g., those which included people referred from secondary care) or which did not recruit only symptomatic patients (e.g., those which included people undergoing population-level screening or referred as a result of screening, polyp surveillance, or with a family history of CRC), were excluded.

A tiered approach to inclusion was taken. Where no or little data for a given test or subgroup was identified, studies that recruited somewhat different populations (e.g., that recruited patients referred from secondary care as well as primary care) were considered for inclusion if generalisability was thought to be reasonable. Where criteria have been widened, this is noted in the report. Decisions around generalisability were made on the basis of the proportion of out-of-scope participants, and on the likely impact of a given patient spectrum. In particular, studies exclusively of screening or surveillance populations were not considered generalisable.

4.1.2 Interventions

Studies were included if they reported data using any of the test-analyser combinations listed in Section 3.3.3, and in Table 1. Data relating to all thresholds were included. Studies reporting dual testing (see Section 3.3.8.4, hereafter referred to as "Dual FIT") were included. Each test was considered individually, but an analysis by test-analyser was not conducted since an assumption of equivalence between devices was considered reasonable by the EAG's clinical advisors for OC-Sensor devices, and there were too few studies to conduct such an analysis for FOB Gold, and the same assumption of equivalence has been made.

4.1.3 Comparators

For the review of clinical efficacy: End-to-end RCT studies that compared one diagnostic strategy to current standard of care (under NG12/DG30, see Figure 1) within England were eligible for inclusion.

For the review of diagnostic test accuracy studies and comparative diagnostic test accuracy studies: studies were included if the reference standard was full colonic imaging via colonoscopy or CTC, or if some patients received other reference standards such as index-test-dependent differential references standards comprising imaging for FIT positive patients and records follow-up for FIT negative patients. This was a change to the published protocol, where a tiered approach was planned, which prioritised studies with 100% colonoscopy or CTC reference standards in the first instance.

It is the EAG's view that all the reference standards available were subject to limitations. Full colonic imaging using colonoscopy or CTC is not 100% accurate,[27, 28] and consequently it may be preferable that studies using this as a reference standard should also perform additional follow-up via medical records to identify missed cases. This was rarely if ever done within the studies identified by this review. This reference standard is also not suitable for some patients, such as those who are elderly/infirm and those with rectal bleeding. Studies using this reference standard may therefore exclude some patients from their analysis, which may reduce generalisability of the findings. Reliance only on records follow-up for some patients may also result in missed diagnoses, e.g., through incomplete record keeping, patients moving away, or dying from another cause before a diagnosis is reached. Records follow-up

may also incorrectly classify some patients as false-negatives where follow-up is long (e.g., in the order of years rather than months) allowing time for cancers that were not present at the time of the index test to have developed. It may also be less sensitive to non-cancer diagnoses, since record keeping for such conditions may be less complete. The recent ACPGBI/BSG review found some numerical differences in diagnostic test accuracy between studies when comparing studies with a full colonic imaging reference standard to those with a differential reference standard comprising a mixture of imaging and records follow-up. However, the difference was not statistically significant.

There is some evidence[29] that most patients with a missed CRC will re-present to primary or emergency care within 6 months of their initial consultation, mitigating some of the concerns with follow-up reference standards. Furthermore, exclusion of studies that did not give all patients full colonic imaging would have largely excluded studies that recruited all patients in primary care (see Section 4.2.1), skewing the patient spectrum away from the population of most interest. As such, all reference standards were eligible for inclusion in the review, and a sensitivity analysis was performed to include only studies with >90% colonoscopy or CTC, as was done in line with the ACPGBI/BSG review.[26]

No adjustment for imperfect reference standards was attempted in the statistical synthesis for this assessment.

Comparative diagnostic test accuracy studies that compared two or more of the tests or test-analyser combinations listed in Section 3.3.3 to each other were included, so long as they included a valid reference standard.

4.1.4 Outcomes

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

- Number of CRC diagnoses
- Number/proportion of CRC diagnoses from urgent referrals
- Stage of detected cancers
- Number/proportion of people identified with other bowel pathologies
- Number/proportion of people with advanced adenomas detected, or detected and treated
- Morbidity including adverse events associated with colonoscopy
- Mortality
- Health-related quality of life
- Anxiety associated with waiting for referral or test results due to diagnostic delays, and further diagnostic workup
- Preference for FIT versus colonoscopy

- Risk of CRC (and IBD and AAs) in relevant subgroups according to FIT threshold
- Test failure rates
- Prognostic implications of false-negative results
- Uptake (completion) of FIT in primary care, to include with respect to cultural, demographic, or socioeconomic factors
- Number/proportion of people referred to secondary care
- Number/proportion of people followed up in primary care
- Duration of validity of negative test (implications for follow-up)
- Number/proportion of urgent (2WW suspected CRC) specialist appointments
- Number/proportion of urgent (2WW suspected CRC) colonoscopy/CTCs
- Number/proportion of non-urgent colonoscopy/CTCs
- Time to colonoscopy/CTC
- Time to diagnosis of CRC or other conditions
- Number/proportion of colonoscopy/CTCs that do not detect CRC
- Number/proportion of colonoscopy/CTCs that do not detect significant bowel pathology
- Number/proportion of people presenting to emergency departments with symptoms of CRC.

For the review of diagnostic test accuracy studies, the following outcomes were eligible for inclusion and extraction:

- Number of true-positives, true-negatives, false-positives, and false-negatives, only where all four statistics were reported or could be calculated for CRC, or for IBD or AAs. IBD and AA data was only extracted from studies that also reported CRC diagnostic test accuracy data.
- Other outcomes as listed for the clinical efficacy studies.

Where an outcome was not identified by the review, and was required by the model, these were subsequently reviewed in the searches for modelling parameters (see Section 5.3.4).

4.1.5 Study design

For the review of end-to-end clinical efficacy studies, RCTs or non-randomised controlled trails were eligible for inclusion.

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

Studies not published in the English language were eligible for inclusion if sufficient data could be extracted from non-English language full-texts, or from an existing English language abstract.

Conference abstracts and non-peer-reviewed reports were eligible if the data were presented in a succinct and accessible manner (e.g., a manuscript prepared for submission to a journal), if sufficient methodological details were reported to allow critical appraisal of the study quality, and if results were reported in sufficient detail. Where there were gaps in the available literature, exclusion criteria for conference abstracts and non-English language papers could be relaxed.

4.1.6 Search strategy

A systematic literature review was undertaken to identify evidence on the intervention (FIT assays) and target condition (CRC), following the guidelines developed by the Centre for Reviews and Dissemination (CRD)[30] for reviews in health care and the Cochrane Handbook for Diagnostic Test Accuracy Reviews.[31]

Searches were run in December 2022 based upon those conducted for the ACPGBI/BSG review (March 2022), which was in turn based upon the searches for DG30 (March 2016). Facets of the searches were limited to either 2022 onwards or 2016 onwards, depending on whether the ACPGBI/BSG (for which searches were done in 2022) or DG30 review (for which searches were done in 2016) had searched that facet. Search strategies used subject headings and free text terms including both generic and product names for the interventions and were optimised for each database. No language restrictions were applied. The search strings are reproduced in full in Appendix 1.

Databases searched included:

- MEDLINE-ALL (via Ovid) including Epub ahead of print, In-Process Citations and Daily Update
- EMBASE (Ovid)
- Cochrane Database of Systematic Reviews (CDSR) (Wiley)
- Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley).

The following additional sources searched in order to identify relevant HTA reports, ongoing reviews and clinical trials (respectively):

- INAHTA (searched 13/12/2022)
- NIHR HTA programme website (searched 13/12/2022)
- PROSPERO (searched 13/12/2022)
- ClinicalTrials.gov (searched 13/12/2022)
- EU Trials Register (searched 13/12/2022)
- WHO ICTRP (searched 13/12/2022).

Retrieved records from all searches were downloaded into EndNote for de-duplication and eligibility screening. Reference lists in included articles and relevant systematic reviews were checked for additional studies. Clinical experts were consulted to ensure that no relevant studies had been missed.

4.1.7 Study selection

Studies were selected for inclusion in the review if they met the inclusion criteria detailed in Sections 4.1.1 to 4.1.5. Titles and abstracts were considered for inclusion against the criteria by one reviewer, with a minimum 10% sample checked by a second reviewer. This was conducted in increments of 100 until 100% sensitivity was achieved, and before the remainder were screened, in order to train both reviewers in implementing the criteria. 100% sensitivity was achieved (all relevant studies were identified by both reviewers) during the second batch of 100 records, though specificity was somewhat lower for both reviewers (both included some additional irrelevant titles), which was dealt with during the full text sift. Full texts were obtained and considered for inclusion by one reviewer, with decisions checked by a second reviewer. Any discrepancies were resolved through discussion.

Where multiple publications relating to the same study were identified, only those with relevant outcome data not published in the others were included. Where there was a crossover in locations and dates of recruitment between two or more studies, the largest was included, unless the other publication(s) reported more thresholds or was a better match for the patient populations of interest (see Section 4.2.2.1), in which case a decision on which to include was based on a consideration of all these factors.

4.1.8 Data extraction strategy

The data extraction form used by ACPGBI/BSG guideline group was used as a basis for a *de novo* data extraction form, which was piloted on three studies and adapted as necessary. Several fields were added including fields relating to study population type (see Section 4.2.2.1) and study and patient characteristics (see Section 3.3.4). Study recruitment dates and locations were extracted to aid an assessment of "crossover" with other studies to avoid double counting of patients. Data relating to diagnostic test accuracy were extracted as the absolute numbers of true positives (TPs), true negatives (TNs), false positives (FPs) and false negatives (FNs) where available, or as sensitivity and specificity which was later transformed into TP, TNs etc as described in Appendix 2. the final list of fields extracted included: first author and date; year of recruitment; location; study name; inclusion and exclusion criteria; population characteristics (age, sex, medications that increase GI bleeding, blood disorders); Test-analyser combination; index test methods; reference standard; N recruited; N missing from analysis; N analysed; outcome (CRC, AA or IBD); N with outcome; threshold; diagnostic accuracy metrics; and any additional outcomes as described in Section 3.3.7.

Data included in the ACPGBI/BSG data extraction form was checked by an EAG reviewer against the original publication and checked for completeness against the inclusion criteria for the review for this assessment (e.g., additional thresholds or subgroups). Additional data was extracted where necessary and checked by a second reviewer. Disagreements were resolved through discussion. Authors were contacted to provide missing data or resolve data ambiguities where of key importance to the review.

Data extractions for a small number of studies were not checked by a second reviewer due to time constraints. These include an update to the Nottingham study[32] (see Section 4.3.2) which was received shortly before the report deadline, and studies included in the reviews of patient preferences and the impact of socioeconomic factors (see Sections 4.3.14.3 and 4.3.14.4).[32-35]

4.1.9 Quality assessment strategy

QUADAS-2[36] was used to assess the quality of the included studies. The scoring scheme is provided in Appendix 3. Scores were assigned by one reviewer and checked by a second, with disagreements resolved through discussion. For the review of comparative diagnostic test accuracy (n=3 included studies), quality assessment using QUADAS-C[37] was planned but due to time constraints was not completed.

4.1.10 Synthesis strategy

4.1.10.1 Narrative synthesis methods

Study and patient characteristics were summarised narratively for all the main analyses. Where there was insufficient data for a statistical synthesis, outcomes were synthesised narratively. Where a statistical synthesis was performed, a narrative synthesis of outcomes was not provided in the interest of brevity.

4.1.10.2 Methods for the meta-analysis of diagnostic test accuracy

Diagnostic accuracy was considered separately for each FIT assay type. For tests where data was available from more than one study, pooled estimates of diagnostic parameters were estimated using the modelling approach described in Jones *et al.*[38] The model accommodates estimates of sensitivity and specificity at more than one explicit diagnostic threshold per study. Pooled estimates are produced at all possible thresholds, even where data for a given threshold has not been reported by an empirical study included in the review. Selected thresholds, based on clinical opinion about the most clinically relevant, are presented in this report. The model is summarised in Appendix 4 and full details are provided in the original publication.

Random effects meta-analysis was used to account for the heterogeneity between studies that is generally expected in diagnostic accuracy studies. Reasons for the heterogeneity in sensitivity and

specificity between studies according to study population type (described in Section 4.2.1.2) and reference standard received (see Section 4.2.1.3) were explored using subgroup analyses.

Summary sensitivity and specificity for each test/fitted model was evaluated based on the mean values of the four sets of study-level random effects ($m_{\mu 1}, m_{\mu 2}, m_{\sigma 1}, m_{\sigma 2}$). As described in Jones *et al*.[38] the summary sensitivity and specificity at any threshold value, C_t, can be calculated as:

$$logit(1 - Specificity(C_t)) = \frac{\left(m_{\mu 1} - \log_e(C_t)\right)}{\exp\left(m_{\sigma 1}\right)}$$

$$logit(Sensitivity(C_t)) = \frac{\left(m_{\mu 2} - \log_e(C_t)\right)}{\exp\left(m_{\sigma 2}\right)} \quad (1)$$

Summary sensitivity and specificity were evaluated for thresholds ranging from 2 (the smallest threshold evaluated in the included studies) to 401 (the largest reported threshold).

Results are displayed as receiver operating characteristic (ROC) plots with summary ROC curves of sensitivity vs 1-specificity. Sensitivity and specificity are also plotted individually against threshold with 95% credible intervals (CrI) for the summary estimates illustrating the range of likely values for average diagnostic accuracy of the synthesised studies. 95% prediction intervals (PrI) are also shown, illustrating the between-study heterogeneity and providing a range of values that might be expected in a future study.

Summary sensitivity and specificity are plotted for the full range thresholds (2 to 401) for all FIT test types. Numerical results and 95% CrIs are presented in tables for selected thresholds, only where the selected thresholds are within the range of values evaluated in the reported studies, to avoid extrapolating beyond the observed data.

Analyses were be conducted in R[39] using the JAGS Markov Chain Monte Carlo (MCMC) sampler and the RJAGS interface package.[40] Convergence to the target posterior distributions was assessed using the Gelman-Rubin statistic[41] for three chains with different initial values. For all analyses, a burn in of 50,000 iterations of the Markov chain was used, with a further 30,000 iterations retained to estimate parameters after thinning by retaining every 10th sample. Model fit penalising for complexity was compared using the Deviance Information Criterion (DIC).[42] Models with lower values of the DIC are preferred. Model fit for all presented analyses is provided in Appendix 5.

4.2 The analysis plan and rationale

The analysis plan was formulated in response to the available data, following the principles set out in the EAG's protocol and taking into consideration the issues outlined in Sections 4.2.1.1 to 4.2.1.3.

4.2.1 Rationale for the analysis plan

4.2.1.1 Impact of specific symptoms on FIT sensitivity and specificity

The EAG heard from clinical advisors that FIT is a better predictor of CRC risk than symptoms alone, and that sensitivity and specificity of FIT may not differ according to the symptoms reported at presentation. The ACPGBI/BSG review[26] showed that the sensitivity was similar in studies recruiting NG12 high/medium-risk patients compared to studies recruiting DG30 low-risk patients (88.7% (95% CI: 84.4, 92.0) and 88.7% (95% CI:78.1, 95.3), respectively), but that the specificity was numerically different (78.5% (95% CI: 73.0, 83.2) and 88.5% (95% CI: 87.1, 89.9) respectively). Since specificity affects estimates of cost-effectiveness, the EAG decided to subgroup studies according to population type to allow exploration of any potential difference.

4.2.1.2 Population types amongst included studies

Missing patients: The population for this appraisal was all patients presenting to primary care with signs and symptoms suggestive of CRC, as listed in NG12 and DG30. A number of studies were encountered that included both NG12 high/medium-risk and DG30 low-risk patients, but only those that reached secondary care (e.g., recruited all on the 2WW). Such studies will be likely to include nearly all NG12 high/medium-risk patients, as all of these should be referred to secondary care as per the pathways outlined in Section 3.2 and Figure 1, but will likely exclude a proportion of DG30 low-risk patients who are not referred and stay in primary care. If the assumption that symptoms do not impact on FIT sensitivity and specificity is incorrect (see Section 4.2.1.1), it would be important to avoid the exclusion of patients who did not make it to secondary care as this would alter the patient spectrum and may bias the estimates of diagnostic test accuracy.

Enrichment with FIT positives: In addition, studies that recruited patients who had reached the 2WW may well include a proportion of DG30 low-risk patients who were referred on the basis of a positive FIT given in primary care before referral (if the region's GPs were using FIT to guide referral according to DG30). Since DG30 low-risk FIT positives (both true positive and false positives) are usually referred, and DG30 low-risk FIT negatives (both false negatives and true negative) are usually not, the patient spectrum will be enriched with DG30 low-risk FIT positive patients whilst excluding most DG30 low-risk FIT negative patients. The exclusion of DG30 low-risk patients whose first FIT was negative is likely to impact on both sensitivity and specificity, and is likely to result in an overestimation of sensitivity (because disproportionately fewer false negatives are included) and an underestimation of

specificity (because disproportionately fewer true negatives are included). A worked example is provided in Appendix 6 to demonstrate this issue. The extent of this bias will depend on the numbers affected by the referral practice and is not known.

Economic model requirements: It was also useful to the model if diagnostic test accuracy data were available for NG12 high/medium-risk and DG30 low-risk patients separately for the following reasons:

- If test accuracy differs according to population, estimates for diagnostic accuracy in the comparator arm would need to come from studies that recruited DG30 patients
- Estimates of the prevalence of CRC in DG30 patients would also be required by the model, as would prevalence for the whole population presenting to primary care (i.e., DG30+NG12 patients)

4.2.1.3 Reference standards

Section 4.1.3 discusses the relative merits of the different reference standards encountered in this review. In summary, all reference standards have limitations, and the restriction to only studies using >90% colonoscopy or CTC would result in the exclusion of most studies which recruited a spectrum of patients closest to being representative of the target population (all patients in primary care) and a greater dependence on studies which may be enriched with FIT positive patients and excludes some of the primary care patients. The worked example in Appendix 6 considers the impact of an imperfect refence standard on estimates of diagnostic test accuracy.

4.2.2 The analysis plan

4.2.2.1 Study categorisation

For the reasons given in Section 4.2.1.2, the studies have been broadly categorised as follows:

- **Population type 1:** Studies closest to being a representative spectrum of all patients presenting to primary care with symptoms of CRC who meet NG12 or DG30 criteria (minus bypass symptoms). This was for studies which recruited the full spectrum of patients, or those with some minor differences in recruitment criteria (wider or narrower than NG12 and DG30), and where a prior FIT result did not influence recruitment.
- **Population type 2:** Studies closest to being a representative spectrum of NG12 high/medium-risk patients. This was for studies that recruited NG12 high/medium-risk patients (minus bypass symptoms). These were often studies that had recruited patients in secondary care who were referred to the 2WW (i.e., population type 4 studies, see below) and had reported a subgroup specifically of NG12 high/medium-risk patients. Because all or nearly all NG12 high/medium-risk patients should be referred to secondary care, studies recruiting in secondary care should recruit most NG12 high/medium-risk patients.

- **Population type 3:** Studies closest to being a representative spectrum of DG30 low-risk patients. This was for studies that recruited a representative spectrum of DG30 low-risk patients. These were likely to have been recruited in primary care, e.g., in areas using FIT in accordance with DG30.
- **Population type 4:** Unclear/likely unrepresentative spectrum. This was for studies that were not population type 1, 2 or 3 studies, or where it was not clear what criteria were used to select patients either for FIT testing or for referral or both. In particular, this included studies that recruited patients in secondary care who were referred to the 2WW, which is likely to be a mix of NG12 high/medium-risk, DG30 low-risk FIT positives (if implemented in primary care at the time of recruitment), and others that GPs have concerns about. It also included studies from countries that did not use NG12 or DG30 and did not state what their criteria were since it would be unclear how representative such a spectrum would be.
    - We note that it could be assumed that studies recruiting patients in the 2WW are likely to be predominantly NG12 high/medium-risk, but we also expect these studies to be enriched with FIT positive patients, as described above. Equally, studies in other countries recruiting patients who have been referred to secondary care are likely to be similar to NG12 high/medium-risk patients, but again the similarity is unknown.

Categorising studies according to population type was difficult. Authors were contacted for more detail where there was uncertainty, but this did not always lead to complete clarity, partly since it was difficult for the authors to tell how well GPs adhered to guidelines about who to give FIT in primary care and/or refer to secondary care.

4.2.2.2 Main, subgroup and sensitivity analyses

As a conservative approach, the EAG considered restricting the analysis to population type 1 studies as the main analysis, but as can be seen from Table ***2***, very few studies recruited a population wide enough to be considered "all patients", and even amongst these the population was often wider or narrower in some way, especially with respect to bypass symptoms (rectal/anal mass or anal ulceration). The EAG therefore included all study types and explored the impact of each through a series of sensitivity analyses.

The following analyses were conducted where >1 study was available to analyse:

Main analysis: Diagnostic test accuracy for CRC for each test individually

- Each test analysed separately, including all study population types 1-4 together

    - Sensitivity analysis removing type 4 studies since these may be enriched with FIT positives, and under the assumption that specific symptoms do not alter FIT sensitivity and specificity
    - Subgroup analysis according to study population type, under the assumption that FIT sensitivity and specificity is affected by specific symptoms:
        - Study population type 1 (all presenting to primary care)
        - Study population type 2 (NG12 high/medium-risk)
        - Study population type 3 (DG30 low-risk)

Additional analysis 1: Diagnostic test accuracy for CRC for all tests together

This analysis was run to allow the investigation of the impact of study population type and reference standards on a larger sample of studies and because these factors were unlikely to interact with test type. It was also used to inform the priors used when less than 5 studies were being synthesised (see Appendix 4).

- All tests analysed together, including all study population types 1-4 together
    - Sensitivity analysis removing type 4 studies
    - Subgroup analysis according to:
        - Study population type 1 (all presenting to primary care)
        - Study population type 2 (NG12 high/medium-risk)
        - Study population type 3 (DG30 low-risk).

This set of studies would also provide estimates of prevalence for each of population types 1-3 for the economic model, since prevalence should not be affected by test type. Similar analyses were planned for diagnostic test accuracy for AA and IBD separately, and undertaken where data allowed.

Additional analysis 2: Impact of reference standard on diagnostic test accuracy estimates

Sensitivity analysis restricting to studies with >90% receiving colonoscopy or CTC as the reference standard, to investigate the effect of the reference standard on estimates of diagnostic test accuracy. This was done for all tests together and tests separately where data allowed.

Table 2: Summary of studies entering the analysis, by test and study population

Test	Main analysis	Population type				Patient characteristics subgroups	Dual FIT	Any analysis
		1: all patients	2: NG12 high/medium-risk	3: DG30 low-risk	4: unclear /unrepresentative			
HM-JACKarc	16	5	4	2	8	Anaemia, sex, age, medications	2	18
OC-Sensor	11	3	1	1	7	Anaemia, sex, medications	1	17
FOB-Gold	3	0	1	0	2	0	0	3
QuikRead go	1	0	1	0	0	0	1	2
NS-Prime	1	0	1	0	0	0	0	1
IDK TurbiFIT	0	0	0	0	0	0	0	0
IDK Hb, Hb/Hp complex	1	0	0	0	1	0	0	1

4.3 Results

The report is structured as follows. The Discussion in Section 6 provides an overview of the evidence base along with a discussion of limitations, and a comparison to other recent reviews. This, along with section 4.3.10, which summarises test accuracy for the tests at selected thresholds, may be a good starting point for understanding the evidence base.

The main analyses for each test are then provided separately in Sections 4.3.1 to 4.3.6. Dual FIT studies are reported in Section 4.3.7. Additional analysis 1 (all tests together, and subgrouped by population type) is provided in Section 4.3.8, and additional analysis 2 (sensitivity analysis for the reference standard) in Section 4.3.9. A summary of the main analyses is provided in Section 4.3.10. A summary of comparative diagnostic test accuracy studies is provided in Section 4.3.11. Separate sections are

provided relating to subgroup analyses according to patient characteristics (see Section 4.3.12), and for studies reporting the AA and IBD studies (Section 4.3.13). Separate sections are also dedicated to studies reporting non-diagnostic test accuracy data, including test failures, uptake and repeat tests (Section 4.3.14.1) time to diagnosis and other outcomes (Section 4.3.14.2), patient acceptability (Section 4.3.14.3) and sociodemographic factors (Section 4.3.14.4).

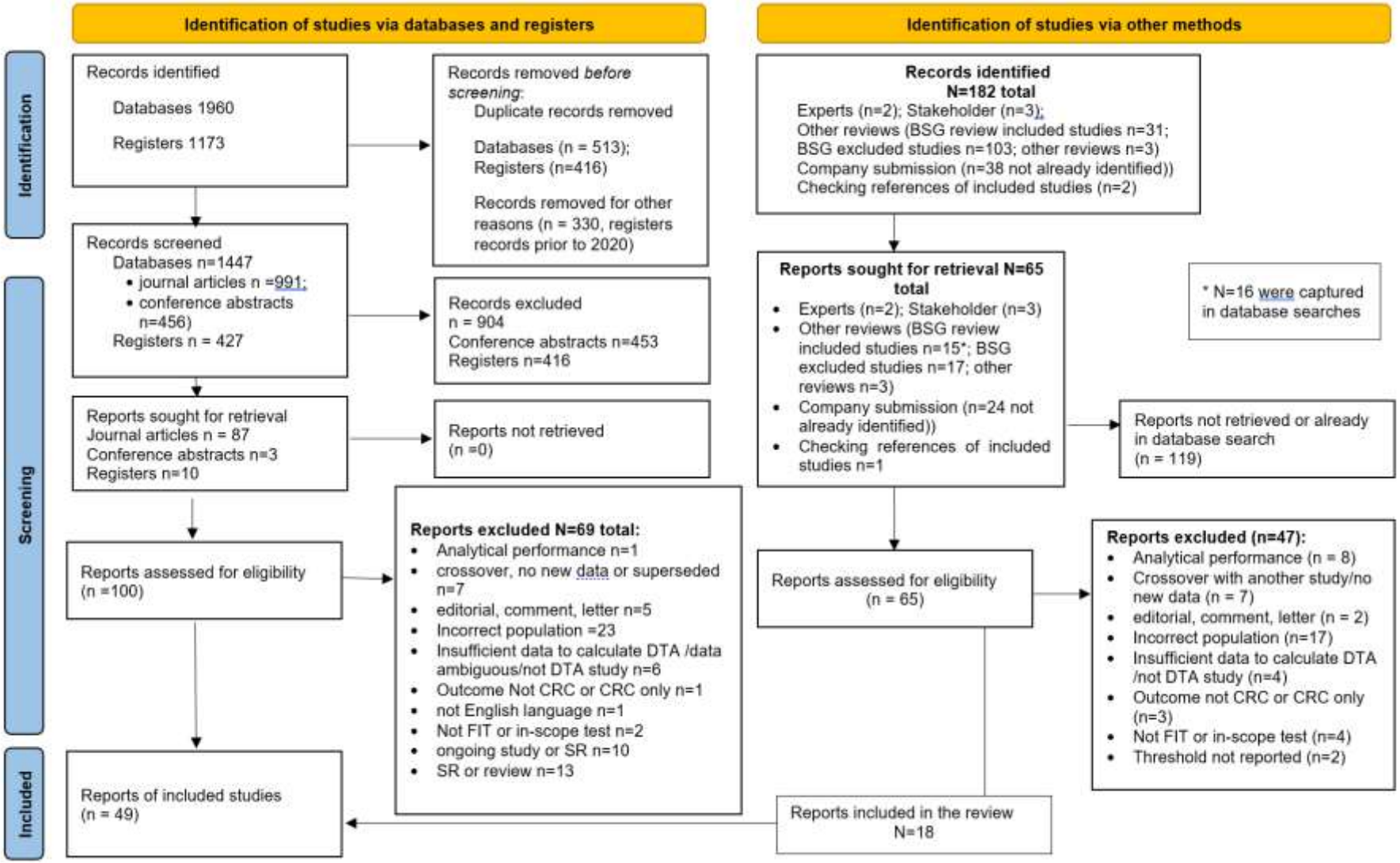
The Prisma flow diagram for the selection of studies is provided in Figure 4. A total of 1874 records were retrieved by the database and registry searches, of which 1774 were excluded on the basis of their title and/or abstract. The full text of the remaining 100 records were retrieved and assessed for eligibility against the study selection criteria. 31 records[29, 34, 35, 43-70] were included in the review. A further 182 records were identified through other sources including nominations by experts or stakeholders (n=5), screening of studies included in other reviews (n=137), company submissions (n=38 not already identified by ScHARR searches) and through the checking of references in other included studies (n=2). From these sources, 18 publications[17, 32, 33, 71-85] were included. In total, 49 publications were included in the review. The records excluded on the basis of their full text are listed in Appendix 7, along with reasons for their exclusion.

No end-to-end studies were identified. Over the 49 included studies, 16 studies reported across 21 publications[17, 29, 48-51, 53, 55, 59, 63, 64, 66, 69, 70, 74-78, 84, 85] reported diagnostic test accuracy data for HM-JACKarc, 17 studies reported across eighteen publications[32, 43-48, 52, 54, 57, 58, 60, 65, 67, 71, 73, 79-81] reported diagnostic test accuracy data for OC-Sensor, three studies reported across three publications[46, 62, 68] reported diagnostic test accuracy data for FOB-Gold, two studies[61, 83] reported diagnostic test accuracy data on QuikRead go, one study[46] reported data for NS-Prime, no studies reported data for IDK TurbiFIT and one[82] reported data for IDK Hb and Hb/Hp complex. Diagnostic test accuracy data relating to patients sub-grouped according to the patient characteristics listed in Section 3.3.4 was identified in 17 publications.[29, 45, 55, 57, 69-71, 73-76, 78-81, 85, 86] Data on dual FIT was identified in 4 publications[54, 78, 83, 84] for HM-JACKarc, OC-Sensor and QuikRead go only. Other outcomes such as test uptake and repeat tests, time to diagnosis, patient preference and sociodemographic factors were reported across 29 publications. [44-49, 51-65, 69, 71, 72, 75, 78, 83-85]

It should be noted that across the evidence base, it was often unclear whether patients with bypass symptoms (rectal or anal mass or anal ulceration) were excluded. Equally, a number of studies excluded patients with rectal bleeding, which may affect the patient spectrum. These issues are not dealt with in detail, due to the poor level of reporting on these factors but should be noted as a potential limitation of the evidence base.

Diagnostic test accuracy data contributing to the analyses are presented in Appendix 8.

Figure 4: Prisma flow diagram of study selection



4.3.1 Main analysis - HM-JACKarc

No end-to-end studies were identified. Seventeen studies reported across 21 publications[17, 29, 48-51, 53, 55, 59, 63, 64, 66, 69, 70, 74-78, 84, 85] reported diagnostic test accuracy data for HM-JACKarc (Table 3). Studies with multiple publications include the NICE FIT study,[17, 75, 76] and a study from Tayside, Scotland, with two publications.[63, 64] Three publications from Oxford[29, 66, 70] comprise a series of different data cuts from a single registry analysis (CSS-BIO-3 4730). These have been counted as two separate studies - one study which recruited January to March 2016,[66] and one study reported over two publications with different but overlapping recruitment dates (recruitment dates March 2017 to March 2020[29] and March 2017 to December 2020[70]) but since both also report unique analyses, both publications were included in the review.

Sixteen studies (17 publications) [29, 48-51, 53, 55, 59, 63, 64, 66, 69, 75, 77, 78, 84, 85] contributed to the main analysis. Patient characteristic subgroup data (see Section 3.3.4) was reported by eight studies,[29, 55, 69, 70, 74, 76, 78, 85] one of which was a study not included in the main analysis because it did not report recruitment dates (so double counting of patients could not be ascertained)[74] and two of which[70, 76] were from studies included in the main analysis (NICE FIT and the Oxford cohort), but the subgroup analysis was reported in a separate publication. Subgroup data is presented in Section 4.3.12. Two studies[78, 84] reported data on dual FIT, one[78] reported data for both single and dual FIT, and one[84] reported data for dual FIT only (see Section 4.3.7). One further study reported data for repeat FITs in Scotland,[56] and another reported a comparison of several different FIT tests[46] (see Section 4.3.11) and was conducted in a subset of the NICE FIT study.[17, 75, 76] One of the NICE FIT publications[17] has been included as it reports AA and IBD data, but has not been used in the analyses relating to CRC as this data is reported in the other NICE FIT publication.[75]

Main analysis: Across the 16 studies (17 publications)[29, 48-51, 53, 55, 59, 63, 64, 66, 69, 75, 77, 78, 84, 85] included in the main analysis, thresholds ranged from 2[17, 49, 63, 64, 75, 85] to 401.[50] Amongst these, an NG12 high/medium-risk subgroup was included from two population type 4 studies[75, 77] since these were likely to be a representative spectrum of NG12 high/medium patients, and to ensure either that the sample was not enriched with FIT positive patients who had received FIT in primary care (NB the NICE FIT study did not include patients who were FIT positive in primary care), and/or because the additional patients were not a full and exclusive spectrum of DG30 low-risk patients. The same was not done to the one Type 4 study[29] which also reported a population type 3 subgroup analysis[70] because the study was not enriched with FIT positives. All studies were in the UK; five (6 publications)[53, 55, 59, 63, 64, 78] were in Scotland and one in Wales[69] with the remainder in England. Sample sizes ranged from 175[51] to 9896[29] and prevalence of CRC from 1.06%[29] to 6.36%.[77] Patient characteristics (sex, ethnicity, blood disorders, medications, anaemia) were rarely or never reported. Age was usually reported as a median, which ranged from 58[66] to 72[50] years amongst studies of types 1-4. The proportion who were male ranged

from 41.4%[56] to 50%.[84] The reference standard was records follow-up in five studies (6 publications),[29, 55, 59, 63, 64, 66] secondary care follow-up comprising various imaging tests in four[50, 51, 84, 85] and colonoscopy or CTC in the remaining seven.[49, 53, 59, 69, 75, 77, 78]

Population type 1 studies: Five studies (six publications)[49, 55, 59, 63, 64, 78] were included in this category. Four studies (five publications)[55, 59, 63, 64, 78] were from Scotland, where the use of FIT has been encouraged in a wider group of patients than in England, encompassing both NG12 high/medium-risk and DG30 low-risk with some differences (see footnotes to Table 3 for Gerrard 2023). CRC prevalence amongst these ranged from 1.29%[59] to 3.05%,[78] possibly indicating heterogeneity in the criteria used to select patients for FIT across these studies, or in how well GPs adhered to guidelines. One further study (D'Souza *et al.* 2020a[49]) was conducted in London at a time when all NG12 and DG30 patients were referred to secondary care by GPs without use of FIT. The prevalence of CRC in this study is higher than the Scottish studies, at 4.03%. The EAG notes that it is likely that not all DG30 patients were referred as GPs would use judgement when making referrals. It reports patients sub-grouped by NG12 high/medium-risk and DG30 low-risk and therefore contributed to three population type subgroup analyses (type 1, type 2, type 3).

Population type 2 studies: Four studies[49, 76, 77, 84] were considered by the EAG to be population type 2 studies because they recruited or reported patients referred to secondary care who met the NG12 high/medium-risk referral criteria. Two of these are subgroups of studies that recruited all patients who were referred to the 2WW (a study from Croydon and the NICE FIT study).[49, 75] All four studies are likely to recruit a fairly representative spectrum of patients meeting NG12 criteria who present to primary care, though where additional criteria, such as a requirement to have undergone a colonscopy[49, 75] was also used to select patients (see Table 3, column 4), some patients may have been systematically excluded (e.g., the elderly). The studies reported thresholds ranging from 2[49, 75] to 150,[75] had sample sizes ranging from 160[49] to 7194[75] and CRC prevalence from 3.57%[75] to 6.36%.[77]

All four had a reference standard that comprised full colonic imaging in secondary care. Two studies[49, 75] did not report whether patients presenting with rectal/anal masses or anal ulceration were included, whilst two others[77, 84] reported small proportions with these symptoms. No data relating to subgroups were reported.

Population type 3 studies: Two studies[49, 70] were considered by the EAG to be population type 3 studies because they reported a subgroup of DG30 patients. Both were subgroups of larger studies.[29, 49] The studies report thresholds 2 and 10µg/g, had samples sizes 138[49] and 166,[70] and CRC prevalence of 1.45%[49] and 0.84%.[70] The reference standard was colonoscopy in one, and records follow-up in another. Neither study reported whether patients presenting with rectal/anal masses or anal ulceration were

included, though as these symptoms are not DG30 criteria, it could be assumed they were not recruited. One study[70] reports multiple patient characteristic subgroups (see Table 3, final column), but for the wider population (type 4).

Population type 4 studies: Eight studies [29, 48, 50, 51, 53, 66, 69, 85] were categorised as population type 4 studies. Four studies[48, 50, 51, 85] included patients referred to secondary care on the 2WW pathway, which may mean the sample is enriched with patients who received FIT in primary care and had a positive test compared to samples not selected on the basis of a positive FIT, e.g., NG12 high risk. Two studies, one from Scotland in 2013[53] and one from Wales in 2020,[69] included patients referred to secondary care using unclear criteria. The two Oxford studies[29, 66] recruited patients given FIT in primary care and are likely to have recruited populations closer to DG30 low-risk type 2 studies, but both included some patients outside of these criteria. In at least one of these, some of the additional patients had symptoms likely to be at lower risk of CRC, including inflammation, thrombocytosis and being tired all the time.

The reference standard was imaging including colonoscopy in all six studies that recruited patients in secondary care,[48, 50, 51, 53, 69] but it was not always clear how many had colonoscopy or CTC. The two studies of patients receiving FIT in primary care[29, 66] used records follow-up as a reference standard. Two studies[69, 85] reported some patients presenting with rectal/anal masses or anal ulceration, one reported 0% of such patients,[29] whilst the remaining studies were unclear for some or all of these criteria. Three studies[29, 69, 85] reported patient characteristic subgroups (see Table 3, final column).

Repeat FIT: one study[56] reported data relating to repeat fit. Data was collected from three regions in Scotland (Tayside, Greater Glasgow & Clyde, Highlands). Patients who returned two FIT tests more than a week apart, but within a year apart were analysed, but it was unclear what criteria were used to select patients for FIT. 10µg/g was the threshold, and records follow-up was the reference standard. The prevalence of CRC was low in this group (0.73%).

Other studies: The other studies reporting patient characteristics subgroup data and Dual FIT data are reported in Sections 4.3.7 and 4.3.12.

Table 3: Study and patient characteristics of HM-JACKarc studies

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
Population type 1 studies (all patient presenting to primary care with symptoms meeting NG12 high/medium or DG30 low-risk)									
1	D'Souza 2020a[49] Croyden, UK Nov 2016 to Oct 2017	HM JACKarc analytical system Colonoscopy	All NG12 and DG30 – all symptomatic patients were referred to colonoscopy in this period in this area of London	NR	mean 60.6 (range 20–90)	• 48.6% • Ethnicity reported[a] • NR	12/298 (4.03%)	2, 10	None
2	Gerrard 2023[78] Lothian, Scotland, UK Jan 2019 to Feb 2020	HM-JACKarc Endoscopy or CT with colorectal protocol.	Urgent suspected of cancer referrals, criteria for referral[a] are both wider and narrower than NG12high/medium and DG30 low-risk	• Wider and narrower than target population[a] • Abdominal mass: 3.0% • rectal mass: 2.4%	Median 65 (IQR 56-74)	• 44.3% • NR • 17.8%	135/3426 (3.05%)	10	Anaemia, no anaemia
3	Johnstone 2022a[55] Greater Glasgow and Clyde, Scotland, UK August 2018 to January 2019	HM-JACKarc (personal communication) Records follow-up	All with NG12 high/medium or DG30 low-risk would get FIT (confirmed by author via personal communication)	May be wider Abdominal mass 2.5%, rectal mass 0.9%	Median 59 (range 16 to 97), n=4968	42.3% NR IDA 5.4%[a]; Anaemia 20.0%	61/4737 (1.29%)	10, 150, 400	Anaemia, no anaemia

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
4	MacDonald 2022[59] NHS Lanarkshire, Scotland, UK October 2016 to February 2019	HM-JACKarc Records follow-up	Symptomatic colorectal referrals from primary care, under SIGN 126 and Scottish Referral Guidelines which encompass both NG high risk and DG30 low-risk for referral	Includes anorectal or abdominal mass; also includes referrals based on imaging, but from GP care.	Median 62 (range 16–96 years)	45.7% NR NR	151/5250 (2.88%)	10	None
5	Mowat 2021[64] & 2019[63] NHS Tayside, Scotland, UK December 2015 to December 2016	HM JACKarc Records follow-up	GPs encouraged to use FIT in patients regardless of the specific lower GI symptoms and perceived risk	NR	Median 65 (range: 2–99, IQR: 51–75)[63]	43.6%[63] NR NR	105/5381 (1.95%)	2, 7, 10, 20, 50, 100, 150, 200, 250, 300, 350, 400	None
Population type 2 studies (NG12 High risk)									
1	D'Souza 2020a[49] Croyden, UK	HM JACKarc analytical system Colonoscopy	NG12 High/medium-risk (subgroup of main Croydon study) who underwent colonoscopy	NR	NR for subgroup		8/160 (5.00%)	2, 10	None

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
	Nov 2016 to Oct 2017								
6	D'Souza 2021a[75] D'Souza 2021c[17d] NICE FIT October 2017 to December 2019	HM JACKarc analytical system Colonoscopy	Subgroup: NG12 High/medium-risk Full study: 2WW patients (including NG12, DG30, others) who underwent colonoscopy	NR	NG12 high/medium-risk: Mean 65.9 (SD 11.1) Full study: 64.0 (SD 11.9)	NG12 high/medium-risk: • 45.7% • Ethnicity reported[a] • IDA 4.2% Full study: • 45.1 • Ethnicity reported[a] • NR	NG12 high risk: 257/7194 (3.57%) Full study: 421/9822 (4.29%)	2, 10, 150	None
7	Farrugia 2020[77] University Hospitals Coventry and Warwickshire NHS Trust, UK January 2015 to March 2017	HM JACKarc automated system Colonoscopy or CT colonography and histology results	NG12 High/med risk[a]	Abdominal/rectal mass n=10	68.6 (error/range NR)	• 48.9% • NR • Anaemia, including iron deficiency 18.1%	10/519 (6.36%)	10	None

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
8	Turvill 2018[84] York Hospital, UK February 2016 to March 2017	HM-JACKarc Full colonoscopy or CT colonography or a lesser investigation (such as CT abdomen/ pelvis with contrast plus flexible sigmoidoscopy) limited by the identification of pathology	NG12 High/medium-risk	4% abdominal mass and 1% rectal mass	median 69 (IQR 61-76)	• 50% • NR • 18% IDA • [a]	27/505 (5.35%)	12	None
Population type 3 (DG30 low-risk)									
1	D'Souza 2020a[49] Croyden, UK Nov 2016 to Oct 2017	HM JACKarc analytical system Colonoscopy	DG30 low-risk (subgroup of main Croydon study) who underwent colonoscopy	NR for subgroup	NR for subgroup		2/138 (1.45%)	2, 10	
9	Withrow 2022[70] (same study as Nicholson 2020)[29a] Oxfordshire, UK	HM JACKarc Records follow-up	**Type 3 subgroup from type 4 study** – FIT given in primary care for any reason, wider than DG30 low-risk alone	NR	Median 61 (IQR 51 to 75)[a]	• 42% • NR • Any anaemia: 26%; IDA: 11%[a]	139/16604 (0.84%)	2, 10	DG30 only subgroup; various anaemia thresholds (men/ women separately); men; women; age </>40,

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
	March 2017 to December 21, 2020 CSS-BIO-3 4730								>50, >60, >70, >80
Population type 4 (unclear/unrepresentative of all presenting to primary care)									
10	Chapman 2021[48] Nottingham University Hospitals Trust, UK Sept 2016 to Sept 2017	HM JACKarc + HM JACKarc analyser 2WW investigations	2WW patients who returned 2 types of FIT test	NR	median 71.1 (IQR 62.5-78.7)		38/732 (5.19%)	4, 10, 22.6, 150	None
11	Elbeltagi 2022[50] North Yorkshire, UK March to October 2020	HM-JACKarc (personal communication) Colonoscopy or cross-sectional imaging	2WW patients	NR	Median72 (IQR: 63-78)	• NR • NR • NR	52/992 (5.24%)	29 thresholds between 6 and 401 at varying intervals	None
12	Faux 2022[51]	HM-JACKarc	2WW patients	Palpable mass 0%	NR	• NR • NR	6/175 (3.43%)	10	None

[d]	**Author, year** **Location** **Recruitment dates** **Study name (if available)**	**Analyser** **Reference standard**	**Inclusion criteria**	**Comparison to NICE scope**	**Mean/median age in years**	**Patient characteristics** • Male; • Ethnicity; Anaemia status	**N with CRC/ N analysed (%)**	**Thresholds µg/g**	**Subgroup data?**
	Cornwall, UK March to July 2020	Colonoscopy or CT abdomen/pelvis, or CT thorax/abdomen/pelvis		anal ulceration NR		• NR			
13	Godber 2016[53] NHS Lanarkshire, Scotland, UK June 2013 to December 2013	HM JACKarc analyser Colonoscopy	Referred to colonoscopy in Scotland, 2013, referral criteria unclear	NR	median 59 (range 16–89), n=507	• 216/507 (42.6%) • NR • 23/484 (4.8%)	11/484 (2.27%)	10	None
9	Nicholson (2018)[66] Oxfordshire, UK Jan to March 2016 CSS-BIO-3 4730	HM JACKarc (NB some pts had two test results, any positive was a positive) Records follow-up	Same criteria as DG30 low-risk, but unknown proportion outside the criteria	NR	Median 58, range 19–93 years	• 43% • NR • n=62 (denominator unclear)	7/238 (2.94%)	7, 10, 20, 50	None
9	Nicholson 2020[29] (overlaps	HM JACKarc Records follow-up	Same criteria as DG30 low-risk, plus some outside the criteria (e.g.,	palpable rectal or anal mass, or anal ulceration n=0	median 60 (range 18-101, IQR 51-74)	• 41.4% • NR • Anameia n=2791/12509	105/9896 (1.06%)	7, 10, 20, 50, 100, 120, 150	Males; females

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
	with Withrow 2022)[70] Oxfordshire, UK March 2017 and March 2020 CSS-BIO-3 4730		Inflammation; thrombocytosis; tired all the time)			= 22.3%; Iron deficiency n=1158/12509= 9.3%			
14	Tang 2022[69] Wales, UK March to June 2020	HM-JACKarc system colonoscopy or CTC, or MPCT (minimal preparation CT)	All consecutive patients referred from primary care on the USC pathway[a]	Abdominal mass 2.5% Anal lump/mass 2.2% Rectal mass 1.5%	median 68 (range 21–97) (n=1050)	• 47.4% • NR • new anaemia 11.1%	20/603 (3.32%)	10	IDA
15	Turvill 2021[85] Yorkshire & Humber, UK April 2018 to Dec 2019 Fast track FIT	HM JACKarc Full colonoscopy or CT colonography, or a lesser investigation (such as CT abdomen/pelvis with contrast or flexible sigmoidoscopy)	2WW patients	Abdominal mass 1.7% Rectal mass 1.6%	Mean 67.3 (SD 11.7) Median 69 (IQR 60, 76)	• 44.5% • NR • IDA or other anaemia, 21.9% • [a]	151/5040 (3.00%)	2	Anaemia, no anaemia, Males, females, medication (antiplatelets, anticoagulants NSAIDs), age </>60
Subgroup data only[c]									

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
16	Cunin 2020[74a] East Sussex, UK NR [must be between 2013 and 2019]	HM JACKarc Various imaging[a]	**Type 2:** NG12 high/medium-risk patients with/without IDA	NR	With IDA: median 74 (IQR 65 to 82) Without IDA: median 72 (IQR 63-79)	With IDA: • 37% • NR • 100% Without IDA: • 41.3% • NR • 0%	With IDA: 20/189 (10.6%) Without IDA: 28/739 (3.79%)	10	IDA, no IDA
6	D'Souza 2021b[76] NICE FIT October 2017 to December 2019	HM JACKarc analytical system Colonoscopy	**Type 4:** NG12 High/medium-risk (subgroup of main NICE FIT study) who underwent colonoscopy	NR	Age <50: 42 (SD6.5) Age 50+: 66.7 (SD9.16)	Age <50: • 40.3% • Ethnicity reported[a] • IDA 5.9%, non-IDA anaemia 1.9% Age 50+ • 45.7% • Ethnicity reported[a] • IDA 4.8%, non-IDA anaemia 5.5%	Age <50: 16/1103 (1.45%) Age 50+: 313/8719 (3.59%)	2, 10, 150	Age </>50
DUAL FIT									
2	Gerrard 2023[78]	HM-JACKarc	**Type 1:** Urgent suspected of cancer referrals,	• Wider and narrower than	Median 65 (IQR 56-74)	• 44.3% • NR • 18.2%	88/2637 (3.34%)	10	0

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
	Lothian, Scotland, UK March 2020 to July 2021	Endoscopy or CT with colorectal protocol.	criteria for referral[a] are both wider and narrower than NG12high/medium and DG30 low-risk	target populationa • Abdominal mass: 3.2% • Rectal mass: 2.4%					
17	Turvill 2018[84] York Hospital, UK February 2016 to March 2017	HM-JACKarc Full colonoscopy or CT colonography or a lesser investigation (such as CT abdomen/ pelvis with contrast plus flexible sigmoidoscopy) limited by the identification of pathology	**Type 3:** NG12 High/med risk	4% abdominal mass and 1% rectal mass	median 69 (IQR 61-76)	• 50% • NR • 18% IDA • Other characteristics [a]	27/476 (5.67%)	43 (either FIT test positive) 2 (both FIT test positive)	0
Repeat FIT									
18	Johnstone 2022b[56] 3 NHS Boards (Tayside, GG&C, Highland), Scotland, UK	HM-JACKarc Records follow-up	**Type 4** since unclear what criteria used to select patients for FIT. Symptomatic patients who had 2 FITs between 1 week and 1 year apart	NR	Median (IQR) GG&C 63 (52-74) Tayside: 69 (56-78) Highland 69 (57-77)	• 41.4% • NR • NR	42/5761 (0.73%)	10	0

[d]	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to NICE scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; Anaemia status	N with CRC/ N analysed (%)	Thresholds µg/g	Subgroup data?
	December 2015 to October 2021[a]								

[a]**Farrugia 2020:** study recruited all 2WW patients, and reported DG30 low-risk and NG12 high/medium-risk subgroups. Only the NG12 subgroup has been included in the analysis since the DG30 subgroup is likely to be highly selected and/or enriched with FIT or Guaiac positive patients; **Gerrard 2023:** Inclusion criteria were urgent suspected of cancer or urgent priority referrals with 'high-risk' symptoms: repeated rectal bleeding without obvious rectal cause or blood mixed in stool, persistent change in bowel habit, palpable abdominal or rectal mass, weight loss and/or abdominal pain with or without unexplained iron deficiency anaemia (IDA); **Johnstone 2022a:** IDA defined as ferritin <15 µg/L; **Johnstone 2022b**: recruitment dates for each area were Tayside: December 2015 to December 2020, Highland: December 2018 to October 2021, Greater Glasgow and Clyde: September 2018 and December 2020; MacDonald 2022: Personal communication with the author indicated that SIGN 126 guidelines and the Scottish Referral Guidelines for Suspected Cancer were used to guide referrals, and that these indicate that both high and low-risk patients as defined by NG12 should be referred. The paper itself lists: "rectal bleeding, diarrhoea, anaemia, anorectal or abdominal mass, abdominal pain weight loss, change in bowel habit (including faecal incontinence), anorectal symptoms (tenesmus, per rectal pain or mucous) or colorectal abnormalities on imaging, per NHSL pre-existing criteria"; **Tang 2022;** referral criteria in Wales unclear; **Turvill 2021:** additional patient characteristic reported, 27% using antiplatelet therapy, anticoagulants or NSAIDs**; Turvill 2018:** additional patient characteristic reported, 30% were taking NSAID, antiplatelet therapy or anticoagulants; **Withrow 2022:** Nicholson 2020 was selected for inclusion in the main analysis despite having fewer patients than Withrow 2022 since it reported more thresholds; **Cunin 2021:** excluded from the main analysis as recruitment dates not reported meaning crossover could not be ascertained, but used in the anaemia subgroup analysis. reference standard was colonoscopy, oesophago-gastroduodenoscopy (OGD), computed tomography (CT) scanning or virtual CT colonography; **D'Souza 2020a**: white (62%), Asian (14%),'other' ethnicity (12%), black (9%) **D'Souza 2021a:** White 5693 (80.0%), Asian 355 (5.0), Black 253 (3.6), Mixed 42 (0.6), Chinese 27 (0.4), Not specified 746 (10.5); **D'Souza 2021b: age <50,** White 68.4%, Asian 9.8%, Black 5.4%, Mixed 1.1%, Chinese 1.2%, Other 18.1%, missing 4.4%

[b] Some type 1-4 studies also report subgroup data as indicated in final column.

[c]The full study population was included in the AA and IBD data analysis as this was not reported for the NG12 subgroup

[d] Study number; GG&C, Greater Glasgow & Clyde

4.3.1.1 Quality assessment

QUADAS 2[36] was applied to studies reporting diagnostic test accuracy data only. QUADAS 2 asks questions relating to internal validity (risk of bias) and external validity (applicability). Each is discussed separately in the following two paragraphs.

Risk of bias: Table 4 summarises the risk or bias and applicability scores as assessed by the authors of this review. Full risk of bias scores are provided in Appendix 3. For risk of bias, no study scored low risk for all items, and no item scored low risk for all studies. The index test scored low risk most often, with only two studies[84, 85] scoring high risk because some or all of the reported thresholds were derived to optimise accuracy. Where patient selection was at risk of bias it was because a consecutive sample was not recruited, and/or because inappropriate exclusions were made, such as excluding people on the basis of not having had a colonoscopy, or not having all blood test results available. Due to the inclusion criteria for the review, all studies avoided a case control design. The reference standard was at unclear or high risk of bias for all studies. This was usually due to not all patients receiving a colonoscopy or CTC, or due to it being unclear if the reference standard was interpreted blind to the index test. Patient flow scored high risk or unclear in nearly all studies. This was due to a mixture of factors, including a lack of clarity about the interval between the index test and the reference standard in nearly all studies, patients receiving different reference standards depending on their FIT result or other factors, and patients being missing from the study.

Applicability: There were concerns about the representativeness of the patients recruited to the studies compared to "all those presenting to primary care" in nearly all studies due to either exclusion of some patients (study population types 2, 3 and 4), or due to a lack of clarity about who was included in comparison to the target population. The index test was at low risk of having poor applicability, except in two cases[29, 66] where a few patients had two index tests and if either scored positive this was counted as a positive test, and two studies[84, 85] scoring high risk because some or all of the reported thresholds were derived to optimise accuracy. The reference standard target condition was CRC in all cases and therefore scored low risk in all studies.

Table 4: HM-JACKarc studies: EAG's assessment of risk of bias and applicability

	Analyses[a]	Risk of Bias items				Applicability items		
		RoB: Patient selection	RoB: Index test	RoB: Reference standard	RoB: Patient flow	Applicability risk: Patients and setting	Applicability risk: Index test	Applicability risk: Reference standard
Benton 2022[46]	2	High	Low	Unclear	Low	High	Low	Low
Chapman 2021[48]	4	High	Low	Unclear	Unclear	High	Low	Low
Cunin 2020[74]	4; anaemia	Low	Low	High	High	High	Low	Low
D'Souza 2020a[49]	1, 2, 3	High	Low	Unclear	Low	Unclear	Low	Low
D'Souza 2021a[75]	2; anaemia	Unclear	Low	Unclear	Unclear	High	Low	Low
D'Souza 2021c[17], D'Souza 2021b[76]	2; age	Unclear	Low	Unclear	Unclear	High	Low	Low
Elbeltagi 2022[50]	4	Unclear	Low	Unclear	High	High	Low	Low
Farrugia 2020[77]	2	High	Low	Unclear	Unclear	High	Low	Low
Faux 2022[51]	4	High	Low	High	High	High	Low	Low
Gerrard 2023[78]	1+/-; single and Dual FIT; anaemia	High	Low	High	High	High	Low	Low
Godber 2016[53]	4	High	Low	Unclear	Low	High	Low	Low

Johnstone 2022a[55]	1; Anaemia	Low	Low	Unclear	High	Low	Low	Low
MacDonald 2022[59]	1+/-	Low	Low	High	High	Unclear	Low	Low
Mowat 2021 & 2019[63, 64]	1	Unclear	Low	High	High	Unclear	Low	Low
Nicholson (2018)[66a]	4	Low	Low	High	High	High	High	Low
Nicholson (2020)[29]	4, sex	Low	Low	High	High	High	High	Low
Tang 2022[69]	4; anaemia	High	Low	Unclear	Unclear	High	Low	Low
Turvill 2021[84]	4; anaemia; age, medications, sex	High	High/Low	High	High	High	High/Low	Low
Turvill 2018[84]	2; single and Dual FIT	High	High	High	High	High	High	Low
Withrow 2022[70b]	3; anaemia; age	High	Low	High	High	High	Low	Low

RoB, Risk of bias

[a] Numbers relate to population-type analyses

[b] Nicholson 2020 and Withrow 2022 include some of the same patients, but it was not clear if the same methodology was used in both studies to select patients and conduct follow-up, so scores are provided for each study based on the information given for that study

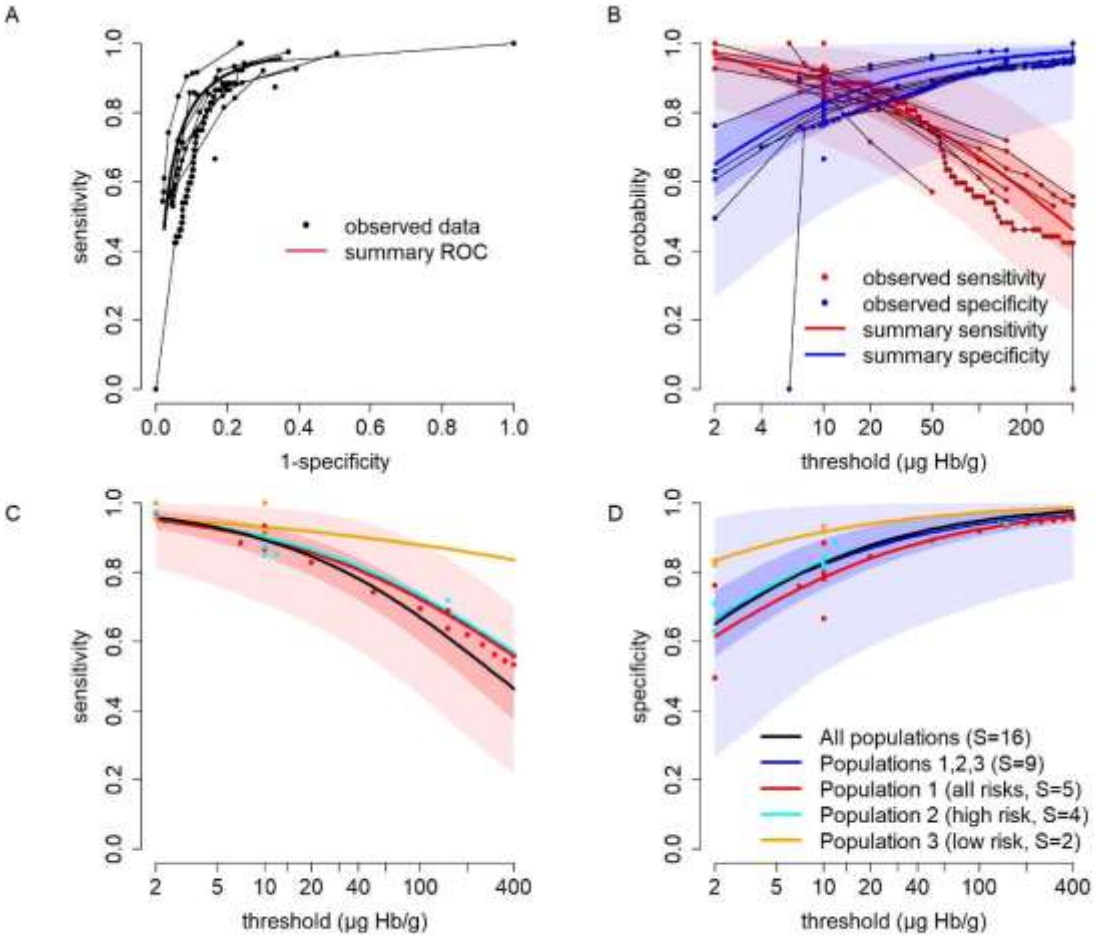
4.3.1.2 Statistical synthesis HM JACKarc

16 studies contributed to the meta-analysis for HM JACKarc. Seven studies provided diagnostic accuracy at a single threshold and the maximum number of thresholds considered within a single study was 103. The final dataset provided a total of 151 pairs of sensitivity and specificity, at thresholds between 2 and 401.

Figure 5 A displays the results on the ROC plane. Observations from the same study are joined. Figure 5 B displays the sensitivity and specificity as a function of threshold. Pooled sensitivity and specificity are shown for subgroups based on population type in Figure 5 C and Figure 5 D, respectively. Sensitivity and specificity for specific thresholds is summarised for all population groups in Table 5.

For the analysis of all studies (populations 1-4), sensitivity ranges from 95.9 (95% CrI: 92.7, 97.9; 95% PrI: 81.4, 99.8) at a threshold of 2, to 46.3 (95% CrI: 37.4, 54.9; 95% PrI: 21.9, 70.2) at a threshold of 400. Specificity ranges from 65.1 (95% CrI: 55.6, 74.8; 95% PrI: 30.3,96.7) at a threshold of 2, to 97.7 (95% CrI: 94.7, 99.2; 95% PrI: 78.1,100). For the analyses of subgroups by population type, the summary estimates were similar and not statistically significant based on overlap of the 95% CrI. The summary sensitivity and specificity for population 3 are higher than for the other considered subgroups, however this analysis was based on only two studies that contributed data at two thresholds (2 and 10). There is therefore considerable uncertainty in the pooled estimates and these should be interpreted with caution.

Figure 5: Observed data and summary sensitivity and specificity for HM JACKarc



95% credible intervals and predictive intervals for summary sensitivity are shown by the dark and light red regions. 95% credible and predictive intervals for summary specificity are shown by the dark and light blue regions.

Table 5: Summary sensitivity and specificity at specific thresholds for HM-JACKarc

Threshold	All studies 1-4 (n=16)		All 1-3 (n=9)		Population 1 (S=5)		Population 2 (S=4)		Population 3 (S=2)	
	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity
2	95.9 (92.7,97.9)	65.1 (55.6,74.8)	95.5 (93,97.1)	66.7 (54.9,77.2)	95.2 (89.3,98.5)	61.6 (41.3,81.8)	95.7 (89.1,98.2)	66.8 (60.4,75.3)	95.5 (83.4,100)	83.3 (74.1,91.2)
2.5	95.3 (91.8,97.5)	68 (58.8,77.3)	95 (92.4,96.7)	69.3 (57.8,79.3)	94.6 (88.4,98.2)	64.2 (44.4,83.9)	95.2 (88.3,97.9)	69.5 (63.2,77.6)	95.3 (83,99.9)	84.8 (75.8,92.3)
3	94.7 (91.1,97.2)	70.3 (61.3,79.3)	94.5 (91.8,96.3)	71.3 (60.2,80.9)	94.1 (87.6,98)	66.3 (46.9,85.4)	94.7 (87.6,97.7)	71.7 (65.4,79.3)	95 (82.6,99.9)	85.9 (77.1,93.1)
4	93.8 (89.8,96.5)	73.7 (65.1,82.2)	93.6 (90.7,95.7)	74.3 (63.7,83.3)	93.3 (86.3,97.6)	69.5 (50.6,87.8)	93.9 (86.4,97.2)	74.8 (68.7,81.9)	94.6 (82,99.9)	87.6 (79.2,94.3)
7	91.4 (86.8,94.8)	79.6 (71.7,87.1)	91.6 (88.3,94.1)	79.5 (70.1,87.2)	91.2 (83.3,96.6)	75.3 (57.1,91.4)	91.9 (83.7,95.9)	80.3 (74.5,86.2)	93.8 (80.6,99.9)	90.3 (82.6,96)
10	89.5 (84.6,93.4)	82.8 (75.2,89.6)	90.1 (86.5,92.8)	82.4 (73.7,89.3)	89.6 (81.1,95.7)	78.6 (60.7,93.2)	90.4 (81.7,94.9)	83.3 (77.8,88.5)	93.2 (79.6,99.8)	91.8 (84.6,96.8)
20	84.7 (79.1,89.6)	87.9 (81.1,93.4)	86.3 (82.1,89.7)	87.1 (79.8,92.6)	85.7 (76.2,93.6)	84.1 (67.1,95.8)	86.7 (77.1,92.1)	88 (83.2,92.2)	NR	NR
50	75.8 (69.4,82)	92.6 (87,96.5)	79.5 (74.5,83.8)	91.7 (86,95.5)	78.8 (68,89.2)	89.5 (74,97.9)	79.9 (69.1,86.7)	92.4 (88.7,95.4)	NR	NR
100	67 (60,74.2)	94.9 (90.3,97.8)	73 (67.1,78.1)	94.1 (89.5,97)	72.2 (60.4,84.7)	92.5 (78.1,98.8)	73.4 (61.3,81.1)	94.7 (91.6,97)	NR	NR
120	64.5 (57.2,71.9)	95.4 (91,98.1)	71 (64.9,76.4)	94.6 (90.3,97.3)	70.2 (58.3,83.3)	93.1 (79.2,98.9)	71.4 (58.9,79.4)	95.2 (92.3,97.4)	NR	NR
150	61.3 (53.7,68.9)	96 (91.9,98.4)	68.5 (62.1,74.2)	95.2 (91.1,97.6)	67.8 (55.5,81.5)	93.8 (80.4,99.1)	68.9 (55.8,77.3)	95.7 (93,97.7)	NR	NR
200	57 (48.9,64.9)	96.6 (92.8,98.7)	65.2 (58.4,71.2)	95.8 (92.1,98)	64.4 (51.7,79.1)	94.6 (81.7,99.3)	NR	NR	NR	NR
400	46.3 (37.4,54.9)	97.7 (94.7,99.2)	56.5 (48.7,63.5)	97.1 (94.1,98.7)	55.8 (41.8,72.6)	96.2 (84.8,99.6)	NR	NR	NR	NR

4.3.2 Main analysis - OC-Sensor

No end-to-end studies were identified. Seventeen studies reported across eighteen publications[32, 43, 45-48, 52, 54, 57, 58, 60, 65, 67, 71, 73, 79-81] reported diagnostic test accuracy data for OC-Sensor (Table 6). One study was reported across two publications.[80, 81]

Amongst the 17 studies, the analyser OC-Sensor iO and PLEDIA were used, and also two analysers not mentioned in the NICE scope (DIANA and MICRO). Clinical advisors to the EAG confirmed with the company that these were calibrated in the same way as the in-scope tests and can be considered equivalent. Twelve studies were in the UK[32, 43, 45-48, 52, 54, 58, 60, 65, 71] (one of which was in Scotland[65]), four (five publications) were in Spain[67, 73, 79-81] and one study was in Denmark.[57] Sample sizes ranged from 120[81] to 37,216[32] and CRC prevalence from 0.59%[45] to 11.65%.[73] Patient characteristics (ethnicity, blood disorders, medications, anaemia) were rarely or never reported. Age was usually reported as a median, which ranged from 61 (IQR 55–77)[47] to 71.1 (IQR 62.5-78.7)[48] years amongst studies of types 1-4, whilst the proportion who were male ranged from 43%[47] to 50%.[43]

Eleven studies contributed to the main analysis[32, 43, 45, 47, 48, 52, 57, 58, 60, 65, 67], three to the population type 1 analysis,[32, 47, 52] one to the population type 2 analysis,[46] one to the population type 3 analysis[45] and seven to the population type 4 analysis. [43, 48, 57, 58, 60, 65, 67] The characteristics of these studies are described in more detail in the following sections. Six studies (7 publications) reported subgroup data,[45, 57, 71, 73, 79-81] and one study reported data on dual FIT.[54] The characteristics of these studies are described in more detail in Sections 4.3.7 and 4.3.12.

No diagnostic test accuracy data was identified that related to OC-Sensor Ceres. The company supplied correlation data between OC-Sensor PLEDIA and OC-Sensor Ceres, and between OC-Sensor iO and OC-Sensor Ceres, conducted in accordance with CLSI EP09-A3 (no reference given) which relates to test bias estimation using patient samples.[87] The number of samples tested was 111, and the R was 0.999 compared to PLEDIA and 0.998 compared to iO. This suggests a high level of correlation between the devices and that measurements are likely to be similar for most samples. However, the ERG note that no details were given relating to the patient population or the methods of the analyses performed, or whether the differences would lead to different clinical decisions at specific thresholds. There was no indication of what a clinically acceptable level of disagreement would be. A formal recommendation of equivalence could not be provided on the basis of the evidence provided.

Main analysis: Eleven studies reported across 11 publications were included in the main analysis.[32, 43, 45, 47, 48, 52, 57, 58, 60, 65, 67] Thresholds ranged from 4[32, 47, 48, 52, 58, 65] to 200µg/g[58] and CRC prevalence from 0.6%[45] to 6.62%.[43] There was one population type 4 study that reported a NG12 high/medium-risk subgroup;[45] the subgroup was included instead of the population type 4 analysis to avoid enriching the

sample with patients who had a positive FIT test in primary care. The Benton *et al*. analysis[46] was not included in the main analysis as recruitment dates and locations cross over with Cama 2022,[47] but this is included in the analysis of population type 2 (NG12 high/medium-risk). The reference standard was records follow-up in 6 studies,[32, 47, 52, 57, 60, 67] whilst the remainder used imaging modalities, but not always 90% CTC or colonoscopy. Most studies did not report whether patients with rectal/anal masses and anal ulceration were excluded, except for the three population type 1 studies and one other that reported 0.3% with palpable masses.[65] Two studies[45, 57] in the main analysis reported patient characteristics subgroup data.

Population type 1 studies (all presenting to primary care): Three studies (all OC-Sensor iO)[32, 47, 52] were considered by the EAG to be population 1 studies because they recruited a population thought to be close to all patients presenting to primary care (see Table 6, column 3). These reported thresholds ranging from 4 to 150µg/g, had sample sizes ranging from 4187,[52] to 37,216[32] and CRC prevalence ranging from 1.39%[47] to 1.74%. All three used records follow-up as the reference standard. Study inclusion criteria were not uniform across studies; all exclude rectal masses, but only one excluded anal ulceration[52] and another abdominal masses.[47] One study also largely excluded IDA[47] and one excluded rectal bleeding.[47] No data relating to subgroups were reported.

Population type 2 studies (NG12 high/medium-risk): One UK study[46] was a population type 2 study (NG12 high risk), which reported thresholds of 1, 10 and 100 using OC -Sensor PLEDIA, had a CRC prevalence of 3.00% and used colonoscopy as the reference standard. This was part of the NICE FIT study, reporting a subgroup of NICE FIT who met NG12 high risk criteria, and who were invited to and also completed 4 FIT tests and colonoscopy (n=233 out of 9822 recruited to NICE FIT). Recruitment dates and location for NICE FIT overlap with a type 1 study,[47] which was preferentially selected for inclusion in the overall OC-Sensor analysis, whilst Benton *et al.*[46] is included in the type 2 subgroup analysis. It was not clear if rectal/abdominal mass and anal ulceration patients were included, and no diagnostic test accuracy data was reported for subgroups according to patient characteristics.

Population type 3 studies (DG30 low-risk): One UK study[45] was a population type 3 study (DG30 low-risk), which reported thresholds from 10 to 150 using OC-Sensor PLEDIA, had a sample size of 2892, a CRC prevalence of 0.6% and used imaging (colonoscopy, CT imaging and colon capsule endoscopy) as the reference standard. It was not clear if rectal/abdominal mass and anal ulceration patients were included. Male and female subgroups were reported.

Population type 4 studies (unclear/unrepresentative of patients presenting to primary care): Seven studies[43, 48, 57, 58, 60, 65, 67] were population type 4 studies, comprising a mixture of studies that recruited only patients referred to the 2WW, or were unclear or unrepresentative of patients in primary

care in some other way (see Table 6, column 3). They used a mix of analysers (see Table 6), reported thresholds from 4[48, 58, 65] to 200,[58] had sample sizes ranging from 116 to 4543, CRC prevalence ranging from 1.56%[57] to 6.62%.[43] Three studies[58, 60, 67] used records follow-up as the reference standard, whilst the remainder used imaging modalities, but not always 90% CTC or colonoscopy. Studies from England[43, 48, 58, 60] recruited patients who were referred to the 2WW since the DG30 update in 2017 and which may therefore have recruited DG30 low-risk patients on the basis of a positive FIT in primary care (see Section 4.2 for discussion of why this is problematic). Studies from elsewhere recruited patients referred to secondary care from primary care. One study[65] included a small proportion (0.3%) of patients with a palpable mass, but otherwise it was unclear if rectal/abdominal mass and anal ulceration patients were included. One study reported a subgroup of IDA/anaemia patients.[57]

Studies reporting subgroup data: In addition to the two studies[45, 57] of types 3 and 4 which reported subgroup data, four additional studies (five publications)[71, 73, 79-81] were included that only reported subgroup data. Two (three publications)[73, 80, 81] of these were included in accordance with the tiered approach to study selection where inclusion criteria were relaxed if evidence was scarce for a given subgroup. Both studies were from Spain, where patients were included who were referred from both primary and secondary care. This may alter the patient spectrum and reduce generalisability to the primary care setting. Prevalence (see Table 6) varied a great deal, reflecting the highly selected nature of some of these cohorts. All studies used colonoscopy as the reference standard. Across all six studies reporting subgroup data, four reported data for ICA/anaemia,[57, 71, 79, 81] two for medications (aspirin users[73] and PPI users[80]) and one for males and females separately.[45] It was not clear if rectal/abdominal mass and anal ulceration patients were included. These studies are also discussed in Section 4.3.12.

Dual FIT using OC-Sensor: One UK type 4 study[54] reported data for Dual FIT (not repeat FIT, see Section 3.3.8.4) at a threshold of 10µg/g, but did not state which OC-Sensor analyser was used. This study recruited both NG12 and DG30 patients, but NG12 patients were only included from June 2020 (recruitment period January 2019 to February 2021). The reference standard was records follow-up. It was not clear if rectal/abdominal mass and anal ulceration patients were included. No subgroup data was reported.

Table 6: Study and patient characteristics of OC-Sensor studies

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to Scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; • Anaemia status[f]	N with CRC/ N analysed (%)	Thresholds, µg/g	Subgroups[e]
Population type 1 studies (all patient presenting to primary care with symptoms meeting NG12 high/medium or DG30 low-risk)									
1	**Crooks 2023**[32] Nottingham, UK Nov 2017 to Nov 2021	• iO • Records follow-up	All referral criteria, except anorectal mass.	Anorectal mass excluded	NR	• NR • NR • NR	514/37216 (1.38%)	4, 10, 20, 40, 100	0
2	**Cama 2022**[47] Hertfordshire, UK June 2019 to Nov 2021	• iO • Records follow-up	All DG30 (low-risk) and most NG12 high risk (see column 6)	IDA, rectal bleeding, rectal or abdominal masses excluded	Median 61 (IQR 55–77)	• 43% (of n=12,231) • NR • 2% IDA; nonIDA 4%	74/5341 (1.39%)	4, 10, 100	0
3	**Georgiou Delisle 2022**[52] Croyden, UK Dec 2019 to Oct 2020	• iO • Records follow up	NICE NG12 and DG30 criteria	Rectal mass or anal ulceration referred straight to 2WW	Mean 65 (range 18–99)	• Male 44.8% • See footnotes* • NR	61/4187 (1.46%)	4, 10, 150	0

Population type 2 studies (NG12 High risk)									
4	**Benton 2022[46]** 50 NHS hospitals across England, UK Oct 2017 to Dec 2019 NICE FIT	• OC Sensor PLEDIA • Colonoscopy	NG12 high risk, who had colonoscopy. Randomised to cohort 1 who were given 4 tests	NR	NR	• NR • NR • NR	7/233 (3.00%)	1, 10, 100	0
Population type 3 (DG30 low-risk)									
5	**Ball 2022[45] (additional data by personal communication)** Sheffield, UK Oct 2019 to Dec 2019	• PLEDIA • Colonoscopy or CT imaging[a] and colon capsule endoscopy	DG30 low-risk[a]	NR	NR for this subgroup	• NR for this subgroup • NR • NR for this subgroup	17/2892 (0.6%)	10, 20, 50, 80, 100, 120, 150	Males; females
Population type 4 (unclear/unrepresentative of all presenting to primary care)									
6	**Archer 2022[43]** Sheffield, UK March 2020 to July 2020	• PLEDIA • CT, colonoscopy	2WW patients	NR	n=514 Mean 64.5 years (SD 12.7 yrs)	n=514 • 50% • NR • n=514 IDA (23%)	11/166 (6.62%)	10, 60, 100	0

7	**Chapman 2021**[48] Nottingham, UK September 2016 to September 2017 Getting FIT	• DIANA • Colonoscopy and additional investigations (e.g., radiology)	2WW patients, returning both FIT tests	NR	median 71.1 (IQR 62.5-78.7)	• 43.9% • NR • NR	38/732 (5.19%)	4, 10, 100	0
8	**Juul 2018**[57] Central Denmark Sept 2015 to Aug 2016 NCT02308384	• DIANA • Records follow-up	Patients with "non-alarm" symptoms of CRC[a]	NR	Mean NR	• 43.9% • See footnote[a] • 12.3%	54/3462 (1.56%)	10	Unexplained anaemia
9	**Laszlo 2021**[58] 24 hospitals and 59 GP practices in UK April 2017 to March 2019	• iO • Colonoscopy 77.7% CTC 14.2% Flexi sig 7.5	2WW patients	NR	Median 67 (range 19–99; IQR 57–75))	• 46.6% • See footnote[a] • 19%	90/3596 (2.50%)	4, 6, 10, 20, 50, 80, 100, 120, 150, 200	0

10	**Maclean 2021a**[60] Royal Surrey NHS Foundation Trust (RSFT), UK End of March 2020 to July 2020	• PLEDIA • Assume records follow-up, as some patients were safety netted	2WW patients	NR	NR	• NR • NR • NR	12/358 (3.35%)	10, 150	0
11	**Mowat 2016**[65] NHS Tayside, Scotland, UK Oct 2013 to March 2014	• iO • Colonoscopy	Symptomatic patients referred from primary care with FIT	Palpable mass 0.3%	Median 64 (range 16–90, IQR 52–73) (n=755)	(n=755) • 45.3% • NR • 9%	28/750 (3.73%)	4, 10	0
12	**Pin Vieito 2021**[67] San Sebastian, Spain Jan 2012 to Dec 2016	• NR • Records follow-up	Patients referred from primary care with FIT	NR	NR	• NR • NR • NR	73/4543 (1.61%)	10, 20	0

Subgroup data only[b]									
13	**Ayling 2019[71]** Derriford Hospital, Plymouth, UK March 2014 to March 2017	• NR • endoscopy or computed tomography scan (NR what type of CT)	Population type 4 - 2WW patients	NR	NR	• NR • For n=428, 99.8% White British • 100% anaemia or IDA	Low Haemoglobin group: 7/178 (3.93%) IDA group: 6/137 (4.38%)	10	IDA; Anaemia
14	**Bujanda 2018[73]** Spain (assume Ourense and San Sebastian)[86] March 2012 to 2014 COLONPREDICT	• NR • colonoscopy	Population type 4 - Symptomatic patients referred from primary and secondary care	NR	Aspirin users: mean 72.5 (SD 9) Aspirin non-users: mean 63.7(SD 14)	• Aspirin users: 58.10%; Aspirin non-users: 48.5% • NR • NR	Aspirin users: 51/485 (10.51%) Aspirin non-users: 299/2567 (11.65%)	20	Aspirin users; Aspirin non-users
15	**Morales-Arraez 2018[79]** La Laguna, Spain April 2016 to Dec 2017	• NR • colonoscopy	Population type 4 - Anaemic patients[d] referred from primary care	NR	Mean 71 (SD 12)	• 33.90% • NR • 100% IDA	28/245 (11.43%)	10	IDA

16	**Rodriguez-Alonso 2018**[80] Barcelona, Spain Sept 2011 to Oct 2012	• MICRO • Colonoscopy	Population type 4 - Symptomatic patients referred from primary and secondary care	NR	PPI users: mean (SE) 64.9 ± 11.3 PPI non-users: mean (SE) 57.3 ± 14.0	• PPI users, 44.20%; non-users, 49.80% • NR • users 18.6%; non-users 5.4%	PPI users: 15/525 (2.86%) PPI non-users: 15/477 (3.14%)	20	PPI use
16	**Rodriguez-Alonso 2020**[81] Barcelona, Spain Sept 2011 to Oct 2012	• MICRO • Colonoscopy	Population type 4 - Symptomatic patients referred from primary and secondary care	NR	NR	• 48.3 • NR • 100% IDA	9/120 (7.5%)	15	IDA
Dual FIT									
17	**Hunt 2022**[54] Lancashire and South Cumbria Cancer Alliance (LSCCA), UK Jan 2019 to Feb 2021	• NR • Records follow-up	Population type 4 - Referred to secondary care [c], returned 2 FITs	NR	Median 66 (range 16–103)	• 44% • NR • NR	317/28622 (1.11%)	10	0

[a] **Ball 2022**: data was also available for all patients on the 2WW, the subgroup of patients who meet NG12 criteria were selected for inclusion in the review. CT imaging was a mix of CTC and other CT imaging modalities; **Juul 2018**: FIT test aimed at those ≥30 years with non-alarm symptoms of CRC, according to GP clinical knowledge and instructions which included: change in bowel habits, abdominal pain, unexplained anaemia, and unspecific symptoms (e.g. fatigue or weight loss), but not for IBS workup. Those aged ≥40 years with rectal bleeding, change in bowel habits >4 weeks, abdominal pain and iron deficiency anaemia recommended to be referred straight to secondary care; **Georgiou Delisle 2022:** White/white British 51.4%, Asian/Asian British 12.6%, Black/Black British 14.8%, Chinese 0.8%, Other 18.6%, mixed 1.5%, Not recorded 0.4%; **Juul 2018**: n (%); Danish: 3280 (94.8), Immigrant Western country: 84 (2.4), Immigrant non-western country: 98 (2.8); **Laszlo 2021**: Black/black British 4.5, Asian/Asian British 6.1, Other Asian 2.0, White 23.5, British mixed 17.9, Multiple/other 5.6, Missing data 40.3

[b] Some type 1-4 studies also report subgroup data as indicated in final column

[c] unrepresentative mix of NG12 high risk and DG30 low-risk due to change in referral criteria part-way through the study

[d] IDA defined as Hb<11.9 g/dL in men and Hb<10.9 g/dL in women, and ferritin ≤ 30 g/dL

[e] 0 – None; 1, IDA or Anaemia; 2, M; 3, F; 4, Ethnicity; 5, Medications which may affect GI bleeding; 6, Blood disorders which may affect the performance of the test; 7, age groups (add age)

[f] Study characteristics relating to medications that may cause GI bleeding and conditions that may affect FIT have been removed, since this data was not reported for any studies.

4.3.2.1 Quality assessment

QUADAS 2[36] was applied to studies reporting diagnostic test accuracy data only. QUADAS 2 asks questions relating to internal validity (risk of bias) and external validity (applicability). Each is discussed separately in the following two paragraphs.

Risk of bias: Table 7 summarises the risk or bias and applicability scores as assessed by the authors of this review. Full risk of bias scores are provided in Appendix 3. For risk of bias, no study scored low risk for all items, and no item scored low risk for all studies. The index test scored low risk most often, with only one study[65] scoring high risk because one of the reported thresholds was selected to maximise sensitivity, and another (two publications) scoring unclear.[80, 81] Where patient selection was at risk of bias it was because it was unclear whether a consecutive sample was recruited, and/or because inappropriate exclusions were made, such as excluding people on the basis of not having had a colonoscopy, or excluding those with rectal bleeding. Due to the inclusion criteria for the review, all studies avoided a case control design. The reference standard was at unclear or high risk of bias for all studies. This was usually due to not all patients receiving a colonoscopy or CTC, or due to it being unclear if the reference standard was interpreted blind to the index test. Patient flow scored high risk or unclear in nearly all studies. This was due to a mixture of factors, including a lack of clarity about the interval between the index test and the reference standard in nearly all studies, patients receiving different reference standards depending on their FIT result or other factors, and patients being missing from the study.

Applicability: There were concerns about the representativeness of the patients recruited to all those presenting to primary care in nearly all studies due to either exclusion of some patients (study population types 2, 3 and 4), or there being a lack of clarity about who was included in comparison to the target population. Some studies[32, 52] were classed as being population type 1 despite scoring poorly for this item since the exclusions were relatively minor (IDA and rectal bleeding), though these limitations should be noted. The index test was at low risk of having poor applicability in all studies. The reference standard target condition was CRC in all cases and therefore scored low risk in all studies.

Table 7: OC-Sensor studies: ScHARR's assessment of risk of bias and applicability

	Analyses[a]	Risk of Bias items				Applicability items		
		Patient selection	Index test	Reference standard	Patient flow	Patients and setting	Index test	Reference standard
Archer 2022[43]	4	High	Low	Unclear	Unclear	High	Low	Low
Ayling 2019[71]	Anaemia	High	Low	Low	High	High	Low	Low
Ball 2022[45] (personal communication)	3, sex	High	Low	High	High	High	Low	Low
Ball 2022[45]	4	Low	Low	High	High	Low	Low	Low
Benton 2022[46]	2	High	Low	Unclear	Low	High	Low	Low
Bujanda 2018[73]	Aspirin	High	Low	Low	Unclear	High	Low	Low
Cama 2022[47]	1	High	Low	High	High	High	Low	Low
Chapman 2021[48]	4	High	Low	Unclear	Unclear	High	Low	Low
Crooks 2023[32], Bailey J 2021a[44]	1	High	Low	High	High	High	Low	Low
Georgiou Delisle 2022[52]	1	High	Low	Unclear	Low	High	Low	Low
Hunt 2022[54]	Dual FIT	High	Low	Unclear	Unclear	High	Low	Low
Juul 2018[57]	4, anaemia	High	Low	High	High	High	Low	Low
Laszlo 2021[58]	4	High	Low	Low	High	High	Low	Low
Maclean 2021a[60]	4	Low	Low	High	High	High	Low	Low
Morales-Arraez 2018[79]	Anaemia	High	Low	Unclear	Unclear	High	Low	Low
Mowat 2016[65]	4	High	High (LoD); Low (10ug/g)	Low	High	High	Low	Low
Pin Vieto 2020[67]	4	Unclear	Low	High	High	Unclear	Low	Low
Rodriguez-Alonso 2018[80], Rodriguez-Alonso 2019[81]	Anaemia, PPIs	High	Unclear	Unclear	Unclear	High	Low	Low

RoB, Risk of bias; LoD, limit of detection

[a] Numbers relate to population-type analyses

4.3.2.2 Statistical synthesis OC-Sensor

Eleven studies contributed to the meta-analysis for OC-Sensor. One study provided diagnostic accuracy at a single threshold and the maximum number of thresholds considered within an individual study was 10 (Laszlo 2021).[58] The final dataset included a total of 44 pairs of sensitivity and specificity estimates, at thresholds between 4 and 200.

Figure 6 A displays the results on the ROC plane. Figure 6 B displays the sensitivity and specificity as a function of threshold. Pooled sensitivity and specificity are shown for subgroups based on population type in Figure 6 C and Figure 6 D, respectively. Sensitivity and specificity for specific thresholds is summarised for all population groups in Table 8.

For the analysis of all studies (populations 1-4), sensitivity ranges from 94.2 (95% CrI: 91.2, 96.7; 95% PrI: 84.6, 99.0) at a threshold of 4, to 54.2 (95% CrI: 48.4, 60.2; 95% PrI: 42.2, 67.2) at a threshold of 200. Specificity ranges from 62.7 (95% CrI: 47.4, 77.2; 95% PrI: 12.0,97.7) at a threshold of 4, to 97.3 (95% CrI: 92.9, 99.3; 95% PrI: 71.9,100) at a threshold of 200. For the analyses of subgroups by population type, the summary estimates were similar and not statistically significant based on overlap of the 95% CrI.

Figure 6: Observed data and summary sensitivity and specificity for OC-Sensor

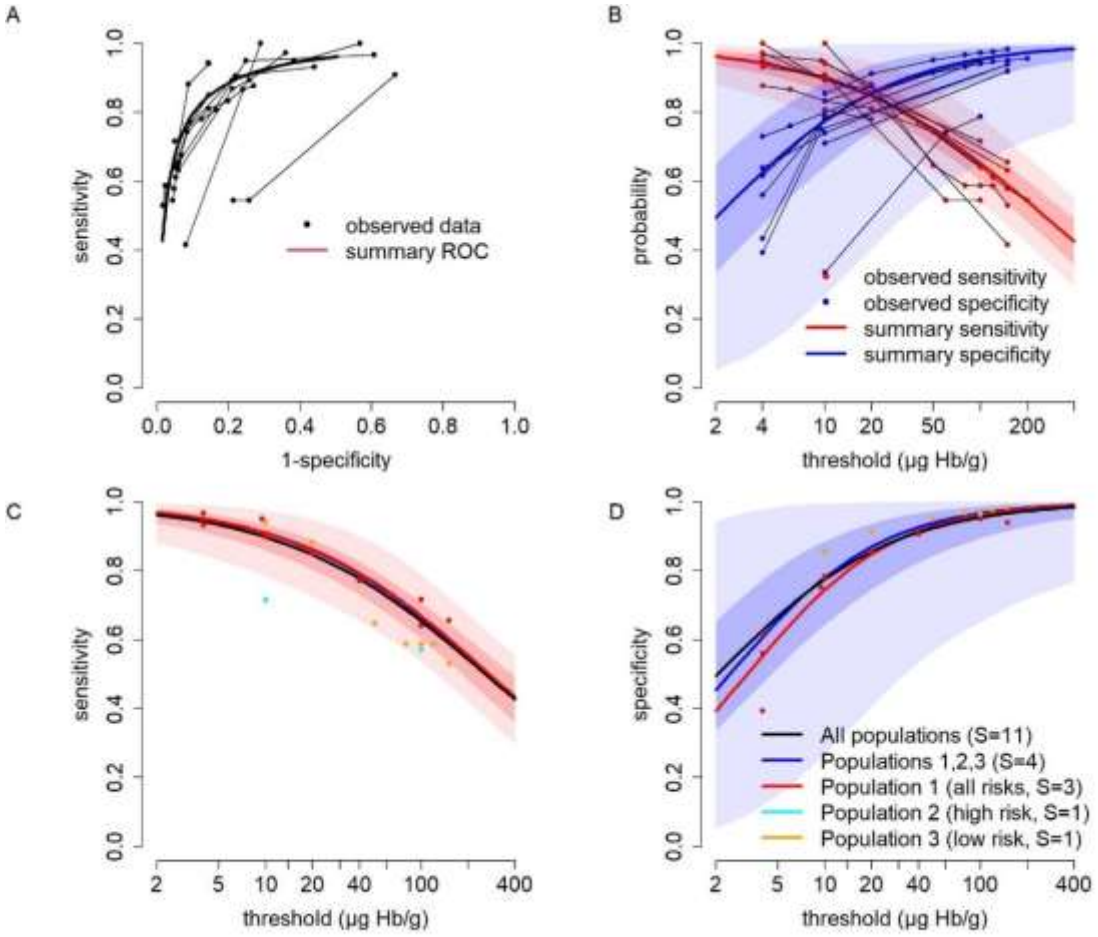


Table 8: Summary sensitivity and specificity at specific thresholds for OC-Sensor

threshold	All studies 1-4 (S=11)		All 1-3 (S=4)		Population 1 (S=3)	
	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity
4	94.2 (91.2,96.7)	62.7 (47.4,77.2)	95 (91.8,97.5)	60.8 (42,77.9)	95.1 (91.1,97.9)	55.3 (36.3,73.5)
7	91.8 (88.2,94.9)	72.3 (58.1,84.8)	92.8 (88.9,96)	72 (54.6,86)	92.9 (88.2,96.7)	67.6 (49.5,83.5)
10	89.8 (85.9,93.3)	77.6 (64.3,88.6)	90.9 (86.6,94.7)	78.1 (62.2,89.8)	91 (85.9,95.5)	74.5 (57.6,88.6)
20	84.7 (80.3,89)	85.6 (74.5,93.6)	86 (80.9,90.9)	87 (74.5,94.7)	86.2 (80.2,92.3)	84.8 (70.5,94.6)
50	75 (70.2,80)	92.5 (84.3,97.3)	76.3 (70.4,82.8)	93.9 (85.8,98)	76.6 (70,84.7)	93 (82.4,98.2)
100	65.3 (60.2,70.7)	95.5 (89.4,98.6)	66.3 (60.2,73.9)	96.6 (91.2,99)	66.8 (60.1,76.2)	96.2 (88.5,99.3)
120	62.5 (57.2,68)	96.1 (90.4,98.9)	63.4 (57.1,71.3)	97.1 (92.3,99.2)	64 (57.2,73.6)	96.8 (89.7,99.4)
150	58.9 (53.4,64.7)	96.7 (91.6,99.1)	59.7 (53.3,67.8)	97.7 (93.4,99.4)	60.3 (53.4,70.2)	97.4 (91.1,99.6)
200	54.2 (48.4,60.2)	97.3 (92.9,99.3)	NR	NR	NR	NR

4.3.3 Main analysis - FOB Gold

No end-to-end studies were identified. Three studies (three publications)[46, 62, 68] reported diagnostic test accuracy data for FOB-Gold (Table *9)*. Two of these studies[46, 62] were comparative diagnostic test accuracy studies which reported data for more than one test-analyser and are also reported in Section 4.3.11. Two studies from the UK[46, 62] used FOB Gold Wide with the SENTiFIT 270 analyser, whilst one study from Norway[68] stated the test to be FOB Gold and the analyser to be Roche Cobas 8000 c702 analyser (Roche Diagnostics, Oslo, Norway). It was not clear if the analysers would produce equivalent data.

No diagnostic test accuracy data was reported for subgroups according to patient characteristics across all three studies, and there was no data on dual FIT using FOB-Gold.

Main analysis: All three studies contributed to the main analysis. Sample size ranged from 163[68] to 553[62] and CRC prevalence from 2.53[62] to 15.95%.[68] Patient characteristics (mean age, sex, ethnicity, blood disorders, medications, anaemia) were not reported except in one case where the proportion who were male was 48.8%.[62] Thresholds ranged from 2[46] to 150µg/g.[62, 68] The reference standard was imaging in all three cases, with two studies recruiting only those who underwent colonoscopy,[46, 68] which may alter the patient spectrum and reduce generalisability, for example since older patients may be offered other imaging modalities such as CTC.

Population type 1 studies (all presenting to primary care): There were no studies of this type.

Population type 2 studies (NG12 high/medium-risk): One UK study,[46] a subgroup of the NICE FIT study,[17, 46, 75, 76] reported a subgroup of patients who met the NG12 high/medium-risk criteria. It reported thresholds 2,10 and 100µg/g using FOB Gold Wide and the SENTiFIT 270 analyser, had a CRC prevalence of 3.00% and used colonoscopy as the reference standard. It was not clear if rectal/abdominal mass and anal ulceration patients were included, nor was the % male reported for this subgroup. No diagnostic test accuracy data was reported for patient characteristic subgroups for the NG12 subgroup, but was available for the wider NICE FIT study using HM-JACKarc (see Section 4.3.1).

Population type 3 studies (DG30 low-risk): There were no studies of this type.

Population type 4 (unclear/unrepresentative of all presenting to primary care): Two studies[62, 68] were unlikely to have recruited a representative sample of the full spectrum of patients presenting to primary care. One study[62] recruited patients referred to the 2WW in England and had a CRC prevalence of 2.53%, whilst the other study[68] recruited patients referred to colonoscopy in Norway and had an exceptionally high prevalence rate of 15.95%. Thresholds ranged from 10 to 150 in both studies, but

Schwettmann *et al*. 2022[68] reported more increments. One used colonoscopy as the reference standard, whilst the other used colonoscopy, CTC or flexible sigmoidoscopy where there were perianal symptoms or anorectal bleeding. It was not clear if rectal/abdominal mass and anal ulceration patients were included. One study[62] had 48.8% male, whilst the other did not report the proportion. No diagnostic test accuracy data was reported for subgroups according to patient characteristics.

4.3.3.1 Quality assessment

QUADAS 2[36] was applied to studies reporting diagnostic test accuracy data only. QUADAS 2 asks questions relating to internal validity (risk of bias) and external validity (applicability). Each is discussed separately in the following two paragraphs.

Risk of bias: Table 10 summarises the risk or bias and applicability scores as assessed by the authors of this review. Full risk of bias scores are provided in Appendix 3. No study scored low risk for all items. The index test conduct scored low risk for all studies. Where patient selection was at risk of bias it was because a consecutive sample was not recruited, and/or because inappropriate exclusions were made, such as excluding people on the basis of not having had a colonoscopy. Due to the inclusion criteria for the review, all studies avoided a case control design. The reference standard was at unclear risk of bias for two[46, 68] studies because it was unclear if the reference standard was interpreted blind to the index test. Patient flow was low risk in one study,[46] high risk in one study[62] due to multiple issues (interval between index test and reference standard unclear, patients receiving a different reference standard on the basis of their FIT result, and patients excluded from the analysis), and one study[68] scored unclear due to a lack of clarity about missing patients.

Applicability: There were concerns about the representativeness of the patients recruited to all those presenting to primary care in all three studies due to either exclusion of some patients (study population types 2 and 4), or there being a lack of clarity about who was included in comparison to the target population. The index test and reference standards were at low risk of having poor applicability in all three studies.

Table 9: Study and patient characteristics of FOB-Gold studies

	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Comparison to Scope	Mean/median age in years	Patient characteristics • Male; • Ethnicity; • Anaemia status	N with CRC/ N analysed (%)	Thresholds, µg/g	Subgroups
Population type 2 studies (NG12 High risk)									
1	Benton 2022[46] 50 NHS hospitals across England, UK Oct 2017 to Dec 2019 NICE FIT	FOB Gold Wide - SENTiFIT 270 Colonoscopy	NG12 high risk, who had colonoscopy. Randomised to cohort 1 who were given 4 tests	NR	NR	• NR • NR • NR	7/233 (3.00%)	2, 10, 100	None

Population type 4 (unclear/unrepresentative of all presenting to primary care)									
2	MacLean 2022a[62] Royal Surrey Foundation Trust, UK July 2019 and March 2020	FOB Gold Wide SENTiFIT 270 Colonoscopy or CTC or flexisig[a]	2WW referrals	NR	NR	• 48.8% • NR • NR	14/553 (2.53%)	10, 100, 150	None
3	Schwettman n 2022[68] Alesund Hospital, Norway January 2020 to February 2021	FOB Gold + Roche Cobas 8000 c702 analyser (Roche Diagnostics, Oslo, Norway) Colonoscopy	Referred to colonoscopy	NR	N	• NR • NR • NR	26/163 (15.95%)	10, 15, 20, 30, 40, 50 ,100, 150	None

[a] **Maclean 2022a:** flexisig if presenting with perianal symptoms or anorectal bleeding

Table 10: FOB-Gold studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores

	Analyses [a]	RoB: Patient selection	RoB: Index test	RoB: Reference standard	RoB: Patient flow	Applicability risk: Patients and setting	Applicability risk: Index test	Applicability risk: Reference standard
Benton 2022[46]	2	High	Low	Unclear	Low	High	Low	Low
MacLean 2022a[62]	2	High	Low	Low	High	High	Low	Low
Schwettmann[68]	4	Unclear	Low	Unclear	Unclear	High	Low	Low

[a] Numbers relate to population-type analyses

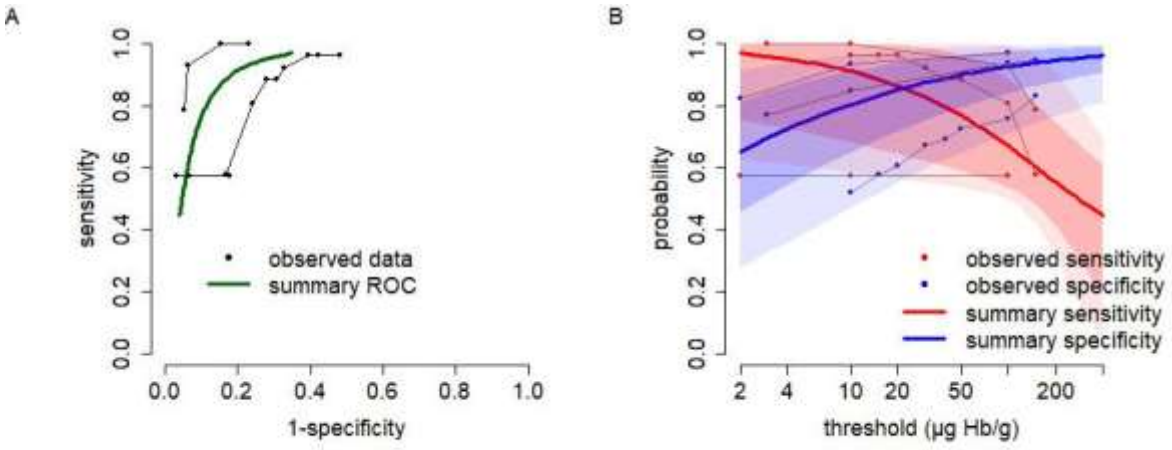
4.3.3.2 Statistical synthesis FOB Gold

Three studies contributed to the meta-analysis for FOB Gold. The number of thresholds considered by each study ranged from 3 to 8 and the final dataset provided a total of 15 pairs of sensitivity and specificity estimates, at thresholds between 2 and 150.

Figure 7 A displays the results on the receiver operating characteristics (ROC) plane. Observations from the same study are joined Figure 7 B displays the sensitivity and specificity as a function of threshold. Due to the small number of studies evaluating FOB Gold subgroup analyses by population type were not conducted. Sensitivity and specificity for specific thresholds is summarised for all population groups in Table 19.

For the analysis of all studies (populations 1-4), sensitivity ranges from 96.9 (95% CrI: 75.6, 100; 95% PrI: 62.5, 100.0) at a threshold of 2, to 60.2 (95% CrI: 48.1, 81.0; 95% PrI: 41.9, 96.2) at a threshold of 150. Specificity ranges from 65.2 (95% CrI: 45.8, 81.1; 95% PrI: 27.9,91.2) at a threshold of 2, to 93.8 (95% CrI: 86.2, 97.9; 95% PrI: 74.2,99.3) at a threshold of 150.

Figure 7: Observed data and summary sensitivity and specificity for FOB Gold



4.3.4 Main analysis - QuikRead go

No end-to-end studies were identified. One study[61] reported diagnostic test accuracy data on QuikRead go that recruited patients exclusively from primary care referrals. Another study[83] reported diagnostic test accuracy data on patients recruited from both primary and secondary care and was included in the analysis of Dual FIT under a tiered approach, since no data meeting the inclusion criteria for the review was identified for this test. It was not included in the analysis of single FIT since data meeting the inclusion criteria were available from MacLean 2022b.[61] Study and patient characteristics of both studies are given in Table 11.

Main analysis: One study[61] was included and analysed. The study was relatively small (n=553) with a small number of CRC events (n=14). It recruited NG12 high/medium-risk patients referred to the 2WW from primary care only. The study was conducted in the UK, had 49.9% males, a CRC prevalence of 2.53%, used colonoscopy, CTC or flexible sigmoidoscopy as the reference standard and reported thresholds of 10, 100, and 150µg/g. It was not clear if patients with rectal/anal mass or anal ulceration were included, and no diagnostic test accuracy data was reported for subgroups according to patient characteristics.

Population type1 studies (all presenting to primary care): There were no studies of this type.

Population type 2 studies (NG12 high/medium-risk): This analysis included the same study as the main analysis.

Population type 3 studies (DG30 low-risk): There were no studies of this type.

Population type 4 (unclear/unrepresentative of all presenting to primary care: There were no studies of this type.

Studies reporting subgroup data: There were no studies of this type.

Dual FIT using QuikRead go: No studies reporting diagnostic test accuracy of dual QuikRead go met the inclusion criteria for the review. One study[83] was included under a tiered approach to inclusion, which reported data for patients referred to colonoscopy from primary and secondary care. It was also relatively small (n=242) with a small number of CRC events (n=13). The study was conducted in Sweden, had 42.1% males, a CRC prevalence of 5.37%, used colonoscopy as the reference standard and reported thresholds of 10, 12 and 20µg/g. It was not clear if patients with rectal/anal mass or anal ulceration were included, and no diagnostic test accuracy data was reported for subgroups according to patient characteristics.

4.3.4.1 Quality assessment

QUADAS 2[36] was applied to studies reporting diagnostic test accuracy data only. QUADAS 2 asks questions relating to internal validity (risk of bias) and external validity (applicability). Each is discussed separately in the following two paragraphs.

Table 11: Study and patient characteristics and diagnostic test accuracy of QuikRead go studies

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Patient characteristics • Mean age in years • Male; • Ethnicity; • Anaemia status	N with CRC/ N analysed (%)	Thresholds, µg/g	Sensitivity (95% CI)	Specificity (95% CI)
Population type 2 studies (NG12 High risk)								
1	MacLean 2021b Royal Surrey Foundation Trust, UK July 2019 and March 2020	QuikRead go Colonoscopy, CTC or flexisig	2WW NG12 high/medium-risk[a]	• Mean age NR • 49.9% • NR • All anaemia: 12.8%; iron or ferritin deficient anaemia: 4.5%	14/553 (2.53%)	10	92.90 (68.5 - 98.7)	70.10 (66.1 - 73.8)
						100	71.40 (45.4 - 88.3)	94.60 (92.4 - 96.2)
						150	57.10 (32.6 - 78.6)	95.90 (93.9 - 97.3)
Dual FIT								
2	Tsapournas 2020 four endoscopy units in Sweden[a] November 2013 to March 2017	QuikRead go Colonoscopy	**Type 4:** Referred for colonoscopy from primary or secondary care	• Median 65 (range 20–87) • 42.1% • NR • NR • Medications[a]	13/242 (5.37%)	See Section 4.3.7		

[a] **Tsapournas 2020:** Eskilstuna General district hospital, Orebro University hospital, Aleris Handen and Hotorget Endoscopy centre, Stockholm. Medications taken by participants reported as Trombyl (aspirin) 23 (9.5), Warfarin 12 (5.0), Others and combinations 8 (3.3); **Maclean 2021b:** population confirmed with author, FIT was not being used by GPs during recruitment period so only NG12 high/medium-risk were referred

Risk of bias: Table 12 summarises the risk or bias and applicability scores as assessed by the authors of this review. Full risk of bias scores are provided in Appendix 3. No study scored low risk for all items. Patient selection was at risk of bias in both studies due to being unclear about or not recruiting a consecutive sample and for excluding patients without a definitive diagnosis or a colonoscopy. Due to the inclusion criteria for the review, all studies avoided a case control design. The index test and reference standards were at low risk of bias in both studies. Patient flow was at high risk due to some patients being missing from the analysis.

Applicability: There were concerns about the representativeness of the patients recruited to all those presenting to primary care both studies due to either exclusion of some patients (study population types 2 and 4; those who did not have a colonoscopy). The index test and reference standard were at low risk of poor applicability in both studies.

Table 12: QuikRead go studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores

	RoB: Patient selection	RoB: Index test	RoB: Reference standard	RoB: Patient flow	Applicability risk: Patients and setting	Applicability risk: Index test	Applicability risk: Reference standard
Maclean 2021b[61]	High	Low	Low	High	High	Low	Low
Tsapournas 2020[83]	High	Low	Low	High	High	Low	Low

RoB, risk of bias

4.3.4.2 Statistical synthesis

Since there was only one study of single FIT and one of Dual FIT, no statistical synthesis was performed. The results of the single FIT study are presented in Table 11.

4.3.5 Main analysis - NS-Prime

No end-to-end studies were identified. Only one study[46] that met the inclusion criteria and reported data for NS-Prime was identified.

Main analysis: One UK study,[46] a subgroup of the NICE FIT study[17, 46, 75, 76], reported a subgroup of patients who met the NG12 high/medium-risk criteria (see Table 13). The study was relatively small (n=233) with a very small number of CRC events (n=7), which may result in less precise estimates. It reported thresholds 3,10 and 100µg/g using NS-Prime, had a CRC prevalence of 3.00% and used colonoscopy as the reference standard. It was not clear if rectal/abdominal mass and anal ulceration patients were included, nor was the % male reported for this subgroup. No diagnostic test accuracy data was reported for patient characteristic subgroups for the NG12 subgroup.

Population type1 studies (all presenting to primary care): There were no studies of this type.

Population type 2 studies (NG12 high/medium-risk): This analysis included the same study as the main analysis.

Population type 3 studies (DG30 low-risk): There were no studies of this type.

Population type 4 (unclear/unrepresentative of all presenting to primary care: There were no studies of this type.

Studies reporting subgroup data: There were no studies of this type.

Dual FIT using NS-Prime: There were no studies of this type.

4.3.5.1 Quality assessment

QUADAS 2[36] was applied to studies reporting diagnostic test accuracy data only. QUADAS 2 asks questions relating to internal validity (risk of bias) and external validity (applicability). Each is discussed separately in the following two paragraphs.

Risk of bias: Table 14 summarises the risk or bias and applicability scores as assessed by the authors of this review. Full risk of bias scores are provided in Appendix 3. Only one study was included in this analysis. It scored low risk for the index test and patient flow items. Patient selection was at risk of bias because a consecutive sample was not recruited, and because patients were asked to complete four tests which may have excluded a spectrum of patients, and only population 2 patients were included. Due to the inclusion criteria for the review, all studies avoided a case control design. The reference standard

Table 13: Study and patient characteristics and diagnostic test accuracy of the NS-Prime study

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Patient characteristics • Mean age in years • Male; • Ethnicity; • Anaemia status	N with CRC/ N analysed (%)	Threshold, µg/g	Sensitivity (95% CI)	Specificity (95% CI)
	Population type 2 studies (NG12 High risk)							
1	Benton 2022[46] 50 NHS hospitals across England, UK Oct 2017 to Dec 2019 NICE FIT	NS-Prime colonoscopy	NG12 high risk, who had colonoscopy. Randomised to cohort 1 who were given 4 tests	• NR • NR • NR • NR	7/233 (3.00%)	3	85.70 (48.7–97.4)	31.90 (26.1–38.2)
						10	71.40 (35.9–91.8)	83.60 (78.2–87.9)
						100	57.1025.1–84.2)	97.30 (94.3–98.8)

95% CI, 95% confidence interval; CRC, colorectal cancer; N, number; NG12 National Guideline 12

Table 14: NS-Prime study: ScHARR's assessment of risk of bias and applicability

	Analyses[a]	RoB: Patient selection	RoB: Index test	RoB: Reference standard	RoB: Patient flow	Applicability risk: Patients and setting	Applicability risk: Index test	Applicability risk: Reference standard
Benton 2022[46]	2	High	Low	Unclear	Low	High	Low	Low

RoB, Risk of Bias

[a] Numbers relate to population-type analyses

was at unclear risk of bias because it was unclear if the reference standard was interpreted blind to the index test.

Applicability: There were concerns about the representativeness of the patients recruited to all those presenting to primary care due to the problems described in the previous paragraph. The index test and reference standards were at low risk of having poor applicability in all three studies.

4.3.5.2 Statistical synthesis

Since there was only one study, no statistical synthesis was performed. The results of the study are presented in Table 13.

. 4.3.7.1Quality assessment

There was insufficient time to synthesise quality assessment for these studies.

4.3.7.2 Statistical synthesis

No statistical synthesis was conducted, since the assessment considered each test individually, and whilst there were two HM-JACKarc studies, there were insufficient threshold points for a meaningful synthesis to be conducted.

4.3.6 Main analysis - IDK tests

IDK TurbiFIT

No end-to-end studies were identified and no diagnostic test accuracy data in patients presenting to primary care with symptoms of CRC were identified for this test. The company submission included an analysis that provided two by two tables for IDK TurbiFIT compared to *IDK* ® Hämoglobin ELISA,[88] but no details were given relating to the patient population, only a simple analysis of agreement between devices in a small sample was given (n=45), and this was only available at one cut off (10µg/g). In this small sample some disagreement between devices in the clinical decisions that would be made at a cut off of 10µg/g was shown, and absolute values were quite different in some samples (e.g. 12.18 compared with 40.51). No assessment of agreement was performed, e.g. using a concordance correlation coefficient or Bland-Altman plot, and no indication of what a clinically acceptable level of disagreement would be was provided. It should also be noted that IDK TurbiFIT and *IDK* ® Hämoglobin ELISA use different test methodologies (immunoturbidimetry and ELISA plates respectively, see Section 3.3.3), and no evidence was provided that linked the *IDK* ® Hämoglobin ELISA or the IDK TurbiFIT to the test used in the clinical study[82] used to support *IDK* ® Hämoglobin ELISA. A formal recommendation of equivalence could not be given on the basis of the evidence provided.

IDK Hb and Hb/Hp tests

No end-to-end studies were identified. Only one study[82] reported diagnostic test accuracy data for IDK Haemoglobin (human) and haemoglobin/haptoglobin complex ELISA tests. This study was conducted in 1999 using a non-commercialised version of this test. IDK have assured the EAG that the data is generalisable to their current test, but it should be noted that no data was offered to support this assertion and the EAG could not validate this statement. There are studies available in screening populations, but these were outside the scope of this assessment since diagnostic test accuracy is expected to differ in asymptomatic populations compared to symptomatic populations.

Main analysis: One German study[46] reported data on symptomatic primary care patients referred to secondary care, but inclusion criteria were otherwise unclear. The study was relatively small (n=621) with a small number of CRC events (n=23). It reported data for both Hb alone and for the complex Hb/Hp. Some data was also reported for Hb + Hb/Hp, but it was not possible to extract TP, TN, FP and FN for this test, nor the sensitivity and specificity. Immunodiagnostik proposed an equation to calculate sensitivity and specificity for the Hb + Hb/Hp test based on the sensitivity and specificity of each test separately. The EAG notes that the use of this equation is usually thought to be valid when the tests are independent, which is not thought to be the case with these two tests, and therefore the EAG has not used this equation. The study had a CRC prevalence of 3.70% and used colonoscopy as the reference standard. It was not clear if rectal/abdominal mass and anal ulceration patients were included, and the % male was 45.1%. No diagnostic test accuracy data was reported for subgroups according to patient characteristics.

Population type 1 studies (all presenting to primary care): There were no studies of this type.

Population type 2 studies (NG12 high/medium-risk): There were no studies of this type.

Population type 3 studies (DG30 low-risk): There were no studies of this type.

Population type 4 (unclear/unrepresentative of all presenting to primary care: This analysis included the same study as the main analysis

Studies reporting subgroup data: There were no studies of this type.

Dual FIT using IDK Hb, Hb/Hp: There were no studies of this type.

4.3.6.1 Quality assessment

QUADAS 2[36] was applied to studies reporting diagnostic test accuracy data only. QUADAS 2 asks questions relating to internal validity (risk of bias) and external validity (applicability). Each is discussed separately in the following two paragraphs.

Risk of bias: Table 15 summarises the risk or bias and applicability scores as assessed by the authors of this review. Full risk of bias scores are provided in Appendix 3. Only one study was included in this analysis, which reported data for two tests, IDK Hb and IDK Hb/Hp complex. It scored high risk for patient selection as it was unclear if patients were recruited consecutively, and only those referred to colonoscopy were included. It scored low risk for the index test for the Hb test, but not for the Hb/Hp complex, since the threshold was derived to optimise accuracy, rather than being prespecified. It scored low risk for the reference standard, but unclear risk for patient flow due to not stating the interval between index test and reference standard, and not stating how many patients were recruited and analysed.

Applicability: There were concerns about the representativeness of the patients recruited to all those presenting to primary care due to the problems described in the previous paragraph. The index test was at high risk for poor applicability because no data relating to the equivalence of this test and the commercial version was presented in symptomatic patients. The reference standard was at low risk of having poor applicability.

Table 15: IDK studies: ScHARR's assessment of risk of bias and applicability

	RoB: Patient selection	**RoB: Index test**	**RoB: Reference standard**	**RoB: Patient flow**	**Applicability risk: Patients and setting**	**Applicability risk: Index test**	**Applicability risk: Reference standard**
Sieg 1999	High	Hb: Low Hb/Hp: High	Low	Unclear	High	High	Low

Table 16*:* **Study and patient characteristics of IDK Hemoglobin (human) and hemoglobin/haptoglobin complex ELISA tests**

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Patient characteristics • Mean age in years • Male; • Ethnicity; • Anaemia status	N with CRC/ N analysed (%)	Thresholds, µg/g	Sensitivity (95% CI)	Specificity (95% CI)
Population type 4 studies (unclear/unrepresentative of patients presenting to primary care):								
1	Sieg 1999[82] Ostringen, Germany NR, prior to publication in 1999	Immunological test for HB Colonoscopy	Referred to secondary care	• Median 59 (range 15-85) • 45.1% • NR • NR	23/621 (3.70%)	2	87.0 (84.4,89.6)	88.1 (85.6,90.6)
		Immunological test for Hb/Hp complex Colonoscopy					82.6 (79.6,85.6)	80.8 (77.7%,83.9)

4.3.6.2 Statistical synthesis

Since there was only one study, no statistical synthesis was performed. The results of the study are presented in Table 16

4.3.7 Main analysis - Dual FIT

No end-to-end studies were identified. Four studies[54, 78, 83, 84] reported data on dual FIT, defined as using two samples from different bowel movements rather than a single sample from one bowel movement to guide referral (see Table 17). This is distinct from the use of a second/repeat FIT during follow-up of patients, i.e., after a decision to refer or not refer has been made.[23] Two studies used HM-JACKarc,[78, 84] one study used OC-Sensor[54] and one study used QuikRead go.[83] Two studies[78, 84] also provided single FIT data for their respective main analyses. One Dual OC-Sensor study[54] only reported data for dual FIT and one QuikRead go study[83] did report data for single FIT, which has been included to aid comparison, but was excluded from the main analysis as the patient cohort included patients referred from secondary care as well as primary care. It is unclear to what extent this might affect generalisability. The characteristics of the studies are presented in full in the respective tables in Sections 4.2.1, 4.2.2 and 4.2.4.

Dual FIT can be interpreted as positive either if both tests are positive ("both" strategy), or if either test is positive ("either" strategy). The "both" strategy is likely to increase specificity, whilst the "either" strategy is likely to increase sensitivity. As the test is being used as a "rule out" test to triage patients to secondary care, it is most useful to maximise sensitivity (the "Sensitive test when Negative rules OUT the disease" (SNOUT) rule),[89] and this is how clinicians indicated the test would be used during the scoping process for this assessment. Therefore, the EAG has concentrated on data relating to the "either positive" interpretation of the test in the synthesis, but presents all data for reference.

One study[54] reported results for both interpretations at 10µg/g, and in this analysis, sensitivity was better in the "either" strategy than in the "both strategy", whilst specificity was higher in the "both" strategy than the "either" strategy. Another study[84] reported the optimal threshold for the "either" and the "both" strategy, showing the same pattern for sensitivity and specificity, but at different thresholds (43µg/g for the "either" strategy and 2µg/g for the "both" strategy).

Two studies[78, 83] reported both dual FIT ("either" strategy) and also data for single FIT. In these studies at 10µg/g sensitivity was better in the dual FIT "either" strategy than single FIT, and specificity was worse, and in the one study[83] that reported multiple thresholds, this general trend continued at higher thresholds.

Table 17: Study and patient characteristics of studies reporting data for dual FIT

		Either positive			Both positive			Single FIT			
Author, year	**N with CRC (or IBD or AA)/N in analysis**	**Thr**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**Thr**	**Sensitivity (95% CI)**	**Specificity (95% CI)**	**N with CRC (or IBD or AA)/N in analysis**	**Thr**	**Sensitivity (95% CI)**	**Specificity (95% CI)**
HM-JACKarc											
Gerrard 2023[78]	CRC: 88/2637 (3.34%)	10	96.60 (90.4-99.3)	71.2 (69.4-73.0)	NR	NR	NR	CRC: 135/3426 (3.94%)	10	93.3 (87.7-96.9)	78.0 (76.6-79.4)
	IBD: 33/2637 (1.39%)		90.90 (75.7 - 98.1)	69.70 (67.9 - 71.5)	NR	NR	NR	55/3426 (1.61%)		90.90 (80.0 - 97.0)	76.30 (74.8 - 77.7)
	AA: 97/2637 (3.68%)		68.00 (57.8-77.1)	70.40 (68.5-72.1)	NR	NR	NR	136/3426 (4.00%)		54.4 (45.6-63.0)	76.40 (45.7-63.0)
Turvill 2018[84]	27/476 (5.67%)	43	87.50 (NR)	90.70 (NR)	2	91.70	85.20	CRC: 27/505 (5.35%)	12	84.60	88.50
OC-Sensor											
Hunt 2022[54]	317/28,622 (1.11%)	10	98 (95.5 to 98.9)	66.20 (65.7 to 66.7)	10	92 (87.9 to 94.1)	81.60 (81.1 to 82.0)	NA	NA	NR	NR
QuikRead go											
Tsapournas 2020[83]	13/242 (5.37%)	10	100.00 (NR)	71.40 (65.5–77.3)		NR	NR	CRC: 13/242 (5.37%)	10	92.30 (77.8–100)	77.30 (71.9–82.7)
		15	92.30 (77.8–100)	76.80 (71.3–82.3)		NR	NR		15	92.30 (77.8–100)	81.70 (76.7–86.7)
		20	92.30 (77.8–100)	81.70 (76.6–86.8)		NR	NR		20	84.60 (65.0–100)	86.50 (82.1–90.9)

Thr, threshold in µg/g; NA, not applicable; NR, not reported

One study[78] reported data at 10µg/g for IBD and AA as well as CRC, in both dual FIT ("either" strategy) and also data for single FIT. Dual FIT had similar sensitivity but lower specificity than single FIT for IBD, and higher sensitivity but lower specificity for AAs, which equated to a 29.8% reduction in missed pathologies.

4.3.8 Additional Analysis 1 - Synthesis of all tests together in a single analysis

This analysis was run to allow the investigation of the impact of study population type and reference standards on a larger sample of studies and because these factors were unlikely to interact with test type. It was also used to inform the priors used when there were less than 5 studies were being synthesised (see Appendix 4).

Twenty-eight studies contributed to the meta-analysis for all tests (OC-Sensor:11, HM JACKarc: 15, FOB Gold: 2). This total is different to a naïve addition of all studies contributing to each test individually because some studies were excluded from the analysis to avoid double counting of patients. This was the case for some studies that reported data for multiple tests and for some studies where the recruitment dates and locations overlapped with other studies.

Eight studies provided diagnostic accuracy at a single threshold and the maximum number of thresholds considered within a single study was 103. The final dataset provided a total of 201 pairs of sensitivity and specificity estimates, at thresholds between 2 and 401.

Figure 8 A displays the results on the ROC plane. Figure 8 B displays the sensitivity and specificity as a function of threshold. Pooled sensitivity and specificity are shown for subgroups based on population type in Figure 8 C and Figure 8 D, respectively. Sensitivity and specificity for specific thresholds is summarised for all population groups in Table 18.

For the analysis of all studies (populations 1-4), sensitivity ranges from 96.5% (95% CrI: 94.8, 97.8; 95% PrI: 86.7, 99.7) at a threshold of 2, to 44.8% (95% CrI: 39.3, 50.0; 95% PrI: 27.0, 61.8) at a threshold of 400. Specificity ranges from 58.7% (95% CrI: 49.9, 67.4; 95% PrI: 15.0,94.9) at a threshold of 2, to 98.1% (95% CrI: 96.3, 99.1; 95% PrI: 80.0,100) at a threshold of 400. For the analyses of subgroups by population type, the summary estimates were similar and not statistically significant based on overlap of the 95% CrI. The summary specificity for population 3 are higher than for the other considered subgroups, however this analysis was based on only three studies (2 HM-JACKarc studies and 1 OC-Sensor). There is therefore considerable uncertainty in the pooled estimates and these should be interpreted with caution.

Figure 8: Observed data and summary sensitivity and specificity for all tests together

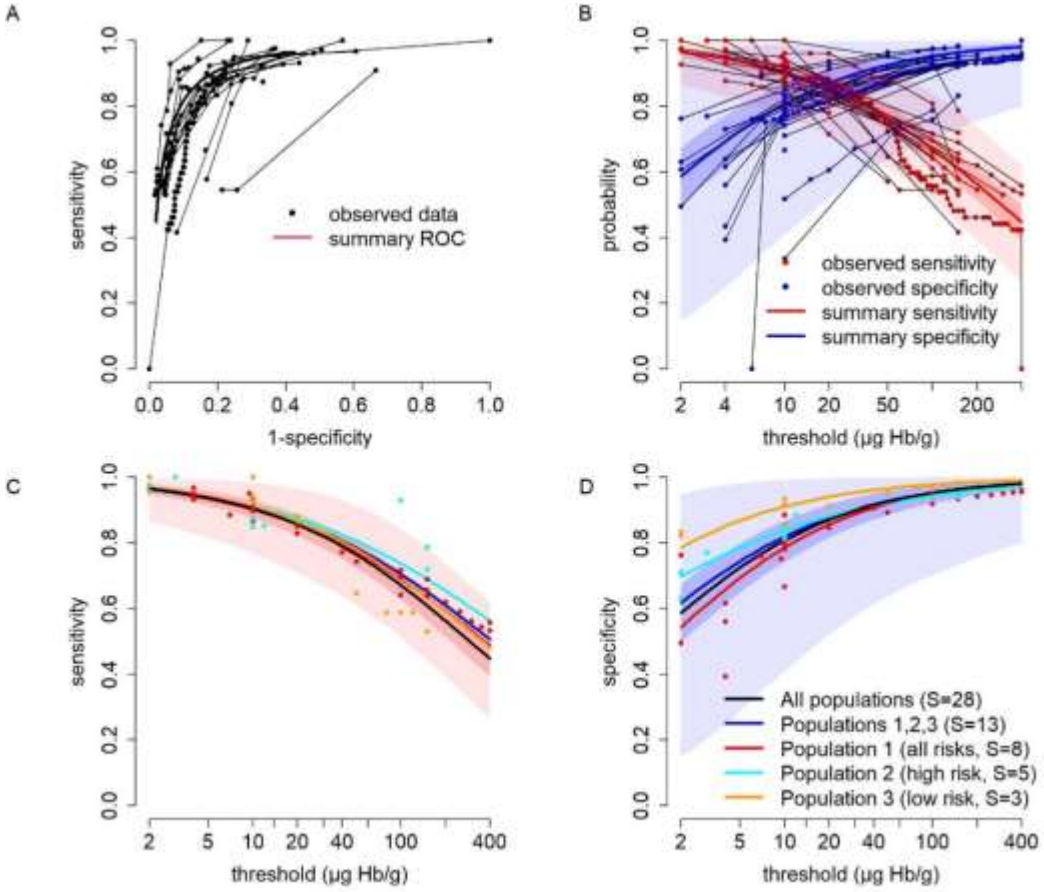


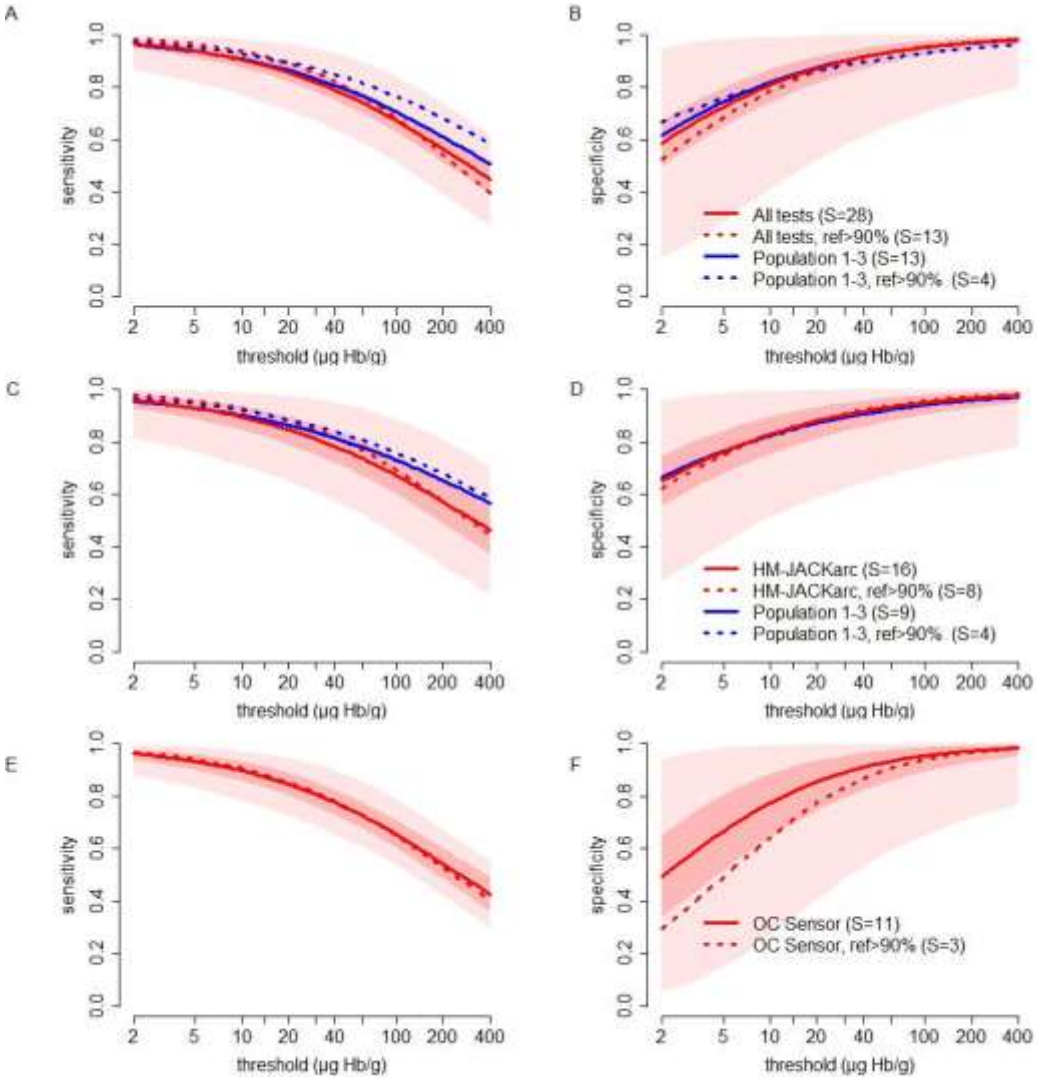
Table 18: Summary sensitivity and specificity at specific thresholds for all tests together

Threshold, µg/g	All studies 1-4 (n=28)		All 1-3 (n=13)		Population 1 (S=8)		Population 2 (S=5)		Population 3 (S=3)	
	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity
2	96.5 (94.8,97.8)	58.7 (49.9,67.4)	96.3 (94.6,97.5)	61.8 (49.2,73.6)	96.2 (93.5,97.9)	54.4 (37.6,71)	96.3 (83.1,99.9)	69.8 (58.9,81.7)	96.7 (89.9,99.6)	78.6 (59.5,93.1)
2.5	96 (94.1,97.4)	62.3 (53.7,70.7)	95.8 (93.9,97.1)	65.1 (52.9,76.4)	95.6 (92.8,97.5)	58.2 (41.6,74.4)	95.8 (82.1,99.9)	72.1 (61.3,83.4)	96.2 (89.1,99.5)	80.9 (62.4,94.3)
3	95.5 (93.5,97.1)	65.1 (56.8,73.3)	95.3 (93.3,96.7)	67.7 (55.8,78.6)	95.1 (92.1,97.2)	61.3 (44.9,77.1)	95.4 (81.1,99.8)	74 (63.3,84.8)	95.7 (88.4,99.4)	82.6 (64.7,95.1)
4	94.6 (92.4,96.4)	69.4 (61.4,77.1)	94.4 (92.3,96.1)	71.6 (60.4,81.8)	94.2 (91,96.6)	65.9 (50.1,80.8)	94.6 (79.6,99.7)	76.7 (66.4,86.7)	94.9 (87.1,99.2)	85.1 (68.1,96.1)
7	92.3 (89.7,94.6)	76.8 (69.7,83.5)	92.4 (89.8,94.4)	78.3 (68.3,87)	92 (88.2,95)	74.1 (59.6,86.9)	92.7 (76.4,99.5)	81.3 (71.7,89.8)	92.8 (84,98.5)	89.1 (73.8,97.7)
10	90.4 (87.6,93)	80.8 (74.3,86.8)	90.7 (87.8,93)	81.9 (72.7,89.7)	90.3 (86.1,93.7)	78.6 (65,89.9)	91.2 (73.9,99.2)	83.9 (74.8,91.5)	91 (81.5,97.8)	91.1 (76.9,98.3)
20	85.6 (82.3,88.8)	87.1 (81.6,91.8)	86.4 (83.1,89.4)	87.6 (79.9,93.6)	85.8 (81,90.3)	85.7 (73.7,94.2)	87.5 (68.6,98.3)	88.1 (79.9,94.1)	86.6 (75.5,95.7)	94.1 (81.9,99.1)
50	76.4 (72.5,80.3)	92.6 (88.5,95.8)	78.5 (74.3,82.4)	92.7 (86.8,96.7)	77.4 (72,83.4)	91.9 (82.4,97.4)	80.5 (59.3,95.8)	92.2 (85.2,96.5)	78 (62.5,90.8)	96.6 (87.1,99.6)
100	67.1 (62.8,71.4)	95.3 (92.1,97.5)	70.5 (65.7,75.3)	95.2 (90.6,98.1)	69 (63.2,76.3)	94.9 (87.2,98.6)	73.7 (50.1,92.2)	94.4 (88.4,97.7)	69.3 (48.4,85.7)	97.8 (90,99.8)
120	64.4 (60,68.7)	95.8 (92.8,97.8)	68.1 (63.2,73.2)	95.7 (91.4,98.3)	66.5 (60.6,74.1)	95.5 (88.3,98.8)	71.6 (47.4,90.9)	94.8 (89.1,97.9)	66.7 (44.3,84.2)	98 (90.7,99.8)
150	60.9 (56.3,65.4)	96.4 (93.6,98.2)	65.1 (59.9,70.5)	96.3 (92.3,98.6)	63.4 (57.3,71.4)	96.1 (89.4,99)	69 (43.4,89.2)	95.4 (89.9,98.1)	63.4 (39.1,82.4)	98.3 (91.5,99.9)
200	56.3 (51.4,61)	97 (94.6,98.5)	61 (55.5,66.8)	96.9 (93.3,98.9)	59.2 (52.8,67.6)	96.8 (90.7,99.3)	NR	NR	NR	98.6 (92.3,99.9)
400	44.8 (39.3,50)	98.1 (96.3,99.1)	50.7 (44.6,57.2)	98 (95.3,99.3)	48.7 (41.7,57.9)	98 (93.4,99.6)	NR	NR	NR	99.1 (94.2,100)

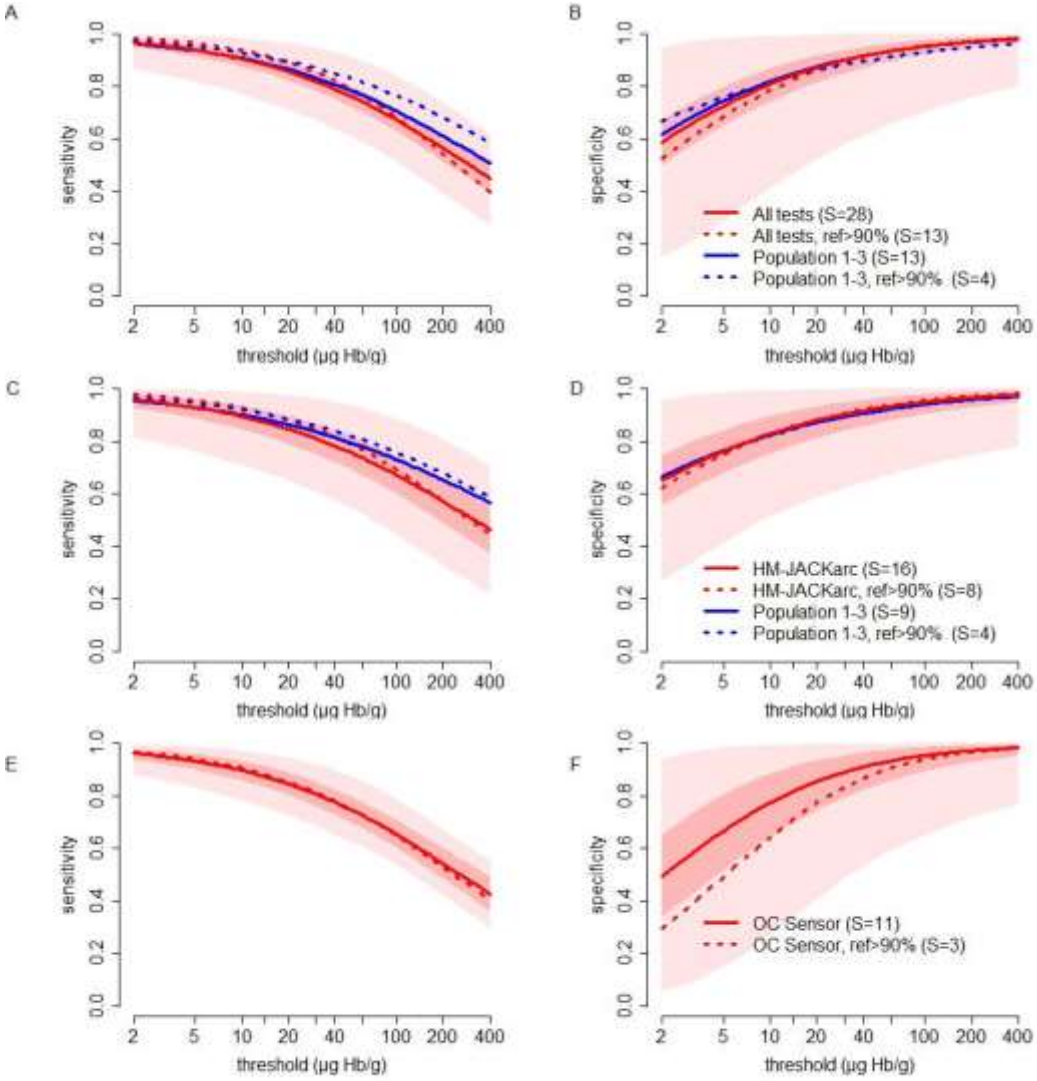
4.3.9 Additional Analysis 2 – Reference standard sensitivity analysis

Subgroup analyses were conducted that included only the studies where at least 90% of the participants received colonoscopy as the reference standard. Subgroups analyses were considered for all FIT tests together, including all population types, excluding population type 4 studies, and separately for each test (where data allowed).

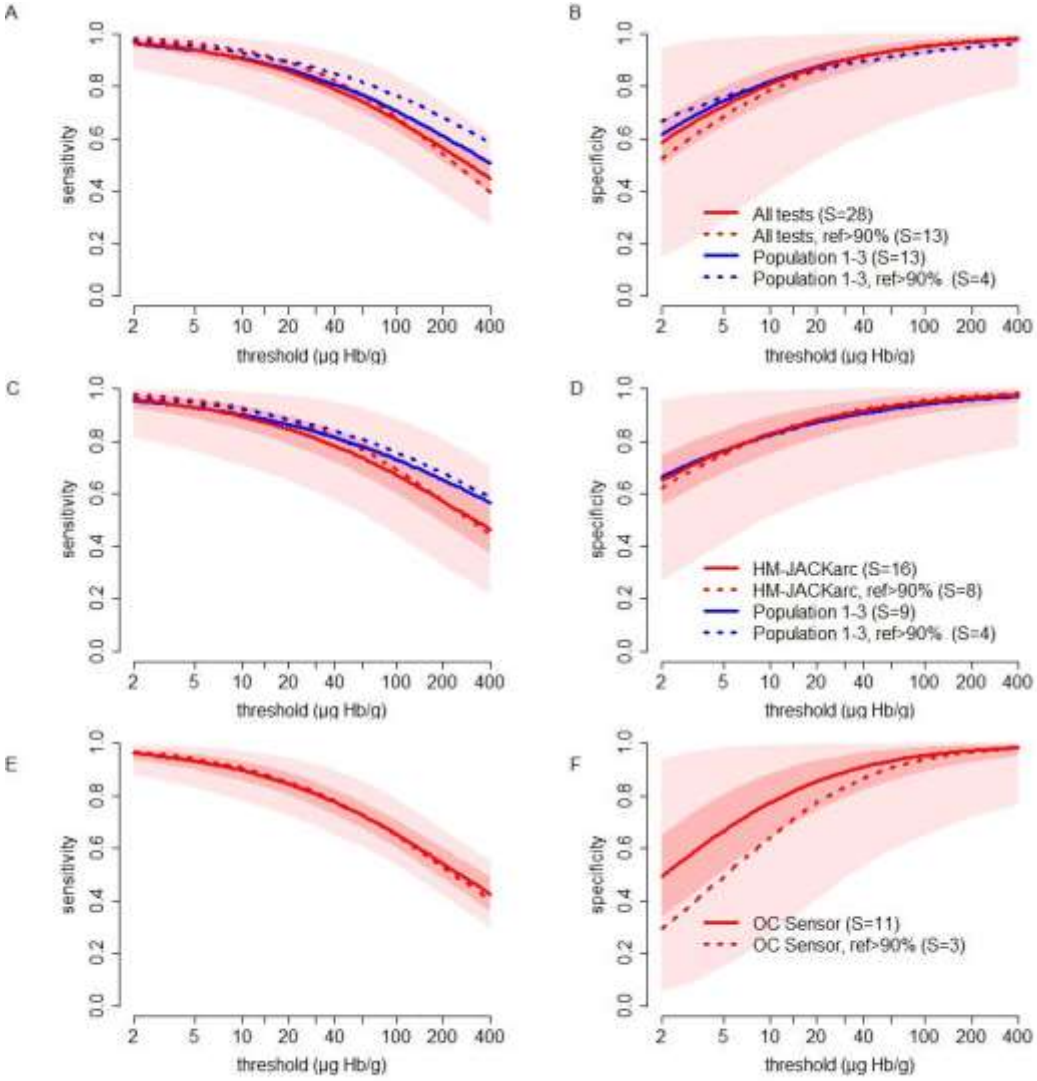
Summary estimates of sensitivity and specificity are illustrated in



. For all analyses the summary estimates were similar, irrespective of the reference standard grouping (all studies vs at least 90% of the participants receiving colonoscopy). The largest difference in point estimates was seen for specificity of OC-Sensor (

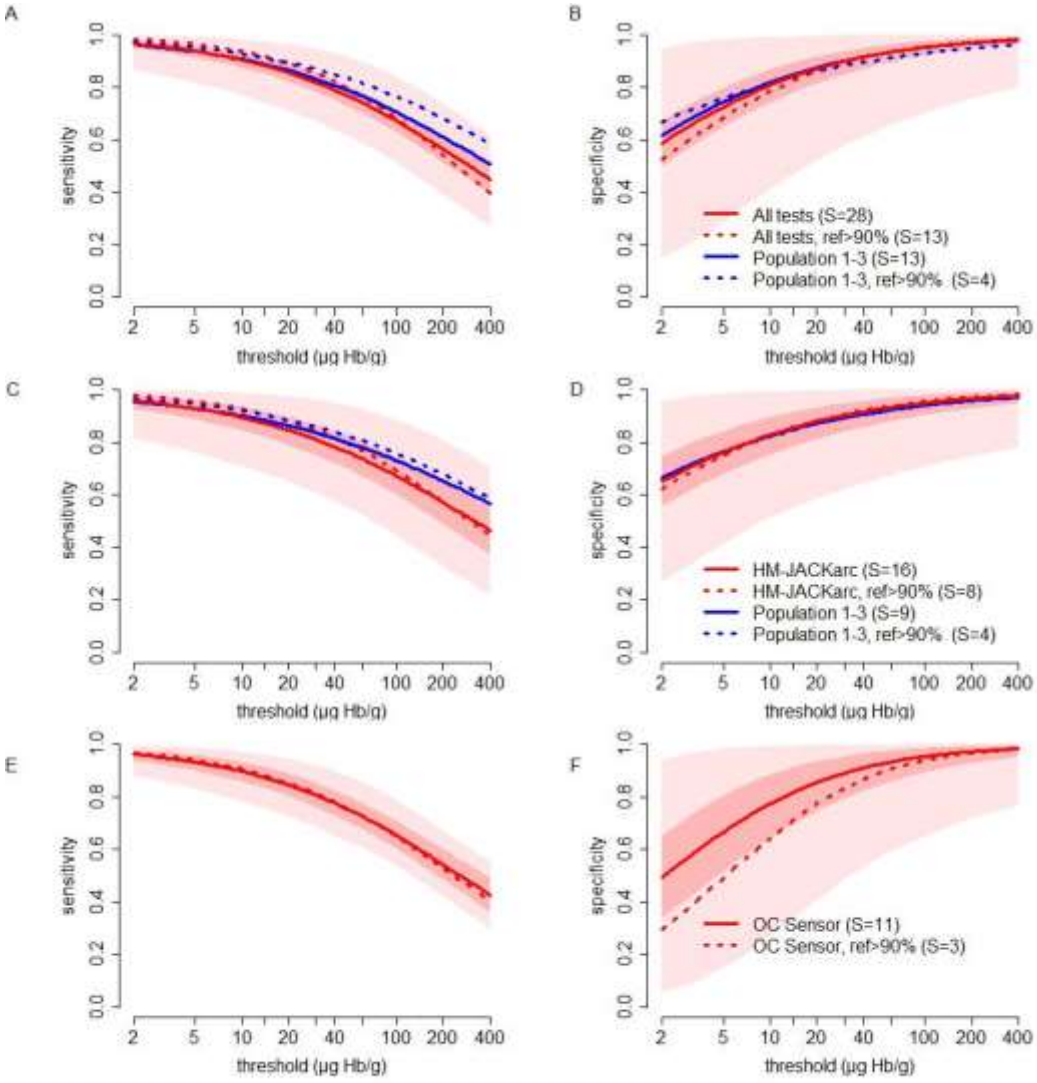


F); however, there were only 3 studies in the >90% colonoscopy subgroup and so the apparent difference may be explained by other sources of heterogeneity between the studies. There was very little difference in specificity for the HM-JACKarc studies -



D).

Figure 9: Summary sensitivity and specificity, reference standard sensitivity analysis



4.3.10 Main analysis - Summary

As described in the previous sections, diagnostic test accuracy was similar across analyses that compared different population classifications and the reference standard used (all studies vs at least 90% of the participants receiving colonoscopy). The analyses used to inform the model are therefore based on all included studies, individually by test type.

The key results for each test type are illustrated in Figure 10 with 95% CrI and PrI shown in dark and light grey, respectively, from the analysis including all FIT test types. Estimated summary sensitivity and specificity at selected thresholds is presented in Table 19 for tests that were statistically synthesised, and also for tests for which there was only one study (NB for these tests, the error is the 95% confidence interval for a single study, rather than the 95% CrI of the summary estimate from the meta-analysis model, and these should not be directly compared). Formal comparison of the different test types was not conducted.

Figure 10: Sensitivity and specificity for all tests

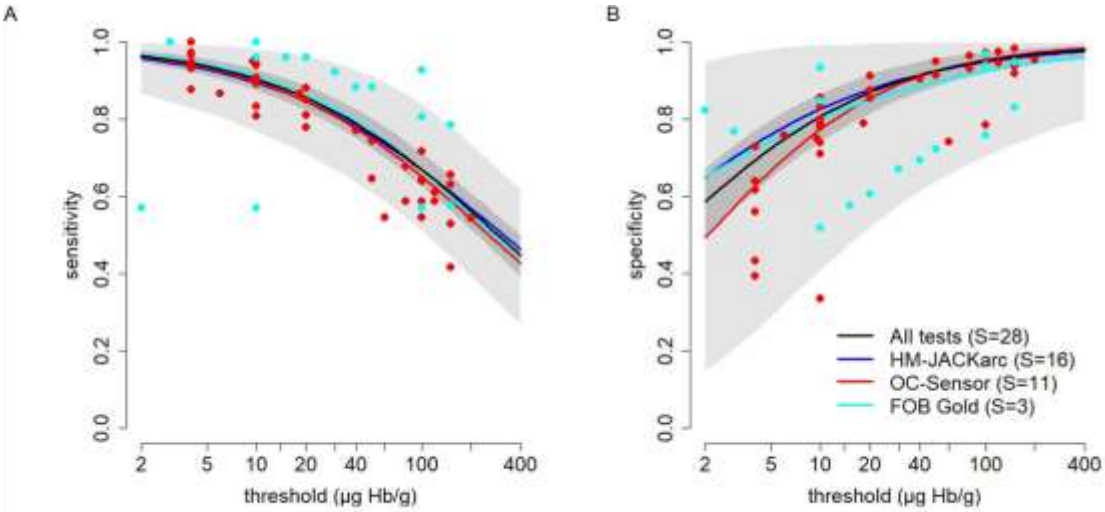


Table 19: Summary sensitivity and specificity at selected thresholds

Threshold (µg/g)	HM-JACKarc (S=16)[a]		OC-Sensor (S=11)[a]		FOB Gold (S=3)[a]		All tests (S=28)[a]		QuikRead go (S=1)[b]		NS-Prime (S=1)[b]		IDK Hb (S=1)[b]		IDK Hb/Hp (S=1)[b]	
	Sens (95% CrI)	Spec (95% CrI)	Sens (95% CrI)	Spec (95% CrI)	Sens (95% CrI)	Spec (95% CrI)	Sens (95% CrI)	Spec (95% CrI)	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)
2	95.9 (92.7, 97.9)	65.1 (55.6, 74.8)	NR	NR	96.9 (75.6, 100)	65.2 (45.8, 81.1)	96.5 (94.8, 97.8)	58.7 (49.9, 67.4)	NR	NR	NR	NR	87 (84.4,89.6)[b]	88.1 (85.6,90.6)[b]	82.6 (79.6,85.6)[b]	80.8 (77.7%,83.9)[b]
2.5	95.3 (91.8, 97.5)	68 (58.8, 77.3)	NR	NR	96.4 (74.7, 100)	67.6 (48.6, 82.9)	96 (94.1, 97.4)	62.3 (53.7, 70.7)	NR	NR	NR	NR	NR	NR	NR	NR
3	94.7 (91.1, 97.2)	70.3 (61.3, 79.3)	NR	NR	96 (73.9, 100)	69.5 (50.8, 84.2)	95.5 (93.5, 97.1)	65.1 (56.8, 73.3)	NR	NR	85.70 (48.7,97.4)[b]	31.90 (26.1,38.2)[b]	NR	NR	NR	NR
4	93.8 (89.8, 96.5)	73.7 (65.1, 82.2)	94.2 (91.2, 96.7)	62.7 (47.4, 77.2)	95.1 (72.6, 100)	72.4 (54.3, 86.2)	94.6 (92.4, 96.4)	69.4 (61.4, 77.1)	NR	NR	NR	NR	NR	NR	NR	NR
7	91.4 (86.8, 94.8)	79.6 (71.7, 87.1)	91.8 (88.2, 94.9)	72.3 (58.1, 84.8)	93 (70, 99.9)	77.5 (60.9, 89.4)	92.3 (89.7, 94.6)	76.8 (69.7, 83.5)	NR	NR	NR	NR	NR	NR	NR	NR
10	89.5 (84.6, 93.4)	82.8 (75.2, 89.6)	89.8 (85.9, 93.3)	77.6 (64.3, 88.6)	91.2 (68.2, 99.8)	80.3 (64.9, 91.1)	90.4 (87.6, 93)	80.8 (74.3, 86.8)	92.90 (68.5 - 98.7)[b]	70.10 (66.1 - 73.8)[b]	71.40 (35.9, 91.8)[b]	83.60 (78.2, 87.9)[b]	NR	NR	NR	NR
20	84.7 (79.1, 89.6)	87.9 (81.1, 93.4)	84.7 (80.3, 89)	85.6 (74.5, 93.6)	86.4 (64.5, 99.4)	85.1 (71.8, 93.7)	85.6 (82.3, 88.8)	87.1 (81.6, 91.8)	NR	NR	NR	NR	NR	NR	NR	NR
50	75.8 (69.4, 82)	92.6 (87, 96.5)	75 (70.2, 80)	92.5 (84.3, 97.3)	76.9 (59.1, 96.4)	89.9 (79.3, 96.1)	76.4 (72.5, 80.3)	92.6 (88.5, 95.8)	NR	NR	NR	NR	NR	NR	NR	NR

Threshold (µg/g)	HM-JACKarc (S=16)[a]		OC-Sensor (S=11)[a]		FOB Gold (S=3)[a]		All tests (S=28)[a]		QuikRead go (S=1)[b]		NS-Prime (S=1)[b]		IDK Hb (S=1)[b]		IDK Hb/Hp (S=1)[b]	
	Sens (95% CrI)	Spec (95% CrI)	Sens (95% CrI)	Spec (95% CrI)	Sens (95% CrI)	Spec (95% CrI)	Sens (95% CrI)	Spec (95% CrI)	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)	Sens (95% CI)	Spec (95% CI)
100	67 (60, 74.2)	94.9 (90.3, 97.8)	65.3 (60.2, 70.7)	95.5 (89.4, 98.6)	67 (53.7, 88.9)	92.6 (83.9, 97.4)	67.1 (62.8, 71.4)	95.3 (92.1, 97.5)	71.40 (45.4 - 88.3)[b]	94.60 (92.4 - 96.2)[b]	57.10 (25.1,84.2)[b]	97.30 (94.3–98.8)[b]	NR	NR	NR	NR
120	64.5 (57.2, 71.9)	95.4 (91, 98.1)	62.5 (57.2, 68)	96.1 (90.4, 98.9)	64 (51.7, 85.5)	93.2 (85, 97.7)	64.4 (60, 68.7)	95.8 (92.8, 97.8)	NR	NR	NR	NR	NR	NR	NR	NR
150	61.3 (53.7, 68.9)	96 (91.9, 98.4)	58.9 (53.4, 64.7)	96.7 (91.6, 99.1)	60.2 (48.1, 81)	93.8 (86.2, 97.9)	60.9 (56.3, 65.4)	96.4 (93.6, 98.2)	57.10 (32.6 - 78.6)[b]	95.90 (93.9 - 97.3)[b]	NR	NR	NR	NR	NR	NR
200	57 (48.9, 64.9)	96.6 (92.8, 98.7)	54.2 (48.4, 60.2)	97.3 (92.9, 99.3)	NR	NR	56.3 (51.4, 61)	97 (94.6, 98.5)	NR	NR	NR	NR	NR	NR	NR	NR
400	46.3 (37.4, 54.9)	97.7 (94.7, 99.2)	NR	NR	NR	NR	44.8 (39.3, 50)	98.1 (96.3, 99.1)	NR	NR	NR	NR	NR	NR	NR	NR

CI, confidence interval; CrI, credible interval; Sens, sensitivity; spec, specificity; S, number of studies

[a] Summary estimates from the meta-analysis model

[b] Individual study estimates. Estimates of error for these studies appear comparatively narrower to those from the synthesis of multiple studies due to being derived from one study only. The number of patients included in each study was: QuikRead go (Type 2 study), n=553, NS-Prime (Type 2 study), n=233, IDK tests (Type 4 study), n=621

4.3.11 Comparative diagnostic test accuracy studies

Three studies conducted a comparison of two or more tests. Chapman *et al*. 2021[48] reported on OC-Sensor DIANA and HM JACKarc, Benton *et al*. 2022[46] compared HM-JACKarc, OC-Sensor PLEDIA, FOB Gold Wide/SENTiFIT 270, and NS-Prime and MacLean *et al*. 2022a compared FOB Gold Wide and QuikRead go. No one test appeared in all three comparisons. Table 20 summarises the study characteristics and reports the sensitivity and specificity at a threshold of 10µg/g, and reports the conclusions drawn by the study authors. The remaining threshold data can be found in the Appendix 8. The largest study included 38 CRC patients amongst a sample of 732.[48] Both other studies[46, 62] had relatively small sample sizes and CRC events (see Table 20).

Different sensitivities and specificities were reported across the tests, and all three studies concluded that at least one test was different to another (see column 10 of Table 20). Due to the small number of CRC events in two of the trials, the small number of studies and the lack of a common comparator, it is difficult to draw any conclusions regarding the comparative performance of the tests, or what and whether different FIT cut-off values are required for each test based on these results. Benton *et al*, who performed an analysis of 4 tests notes that more work is required to understand the clinical impact of the use of different tests.

Table 20: Sensitivity and specificity reported in studies comparing different tests within the same patients

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	N with CRC/ N analysed (%)	Thresholds, µg/g	Threshold	Sensitivity (95% CI)	Specificity (95% CI)	Conclusion drawn by study authors
1	Chapman 2021[48] Nottingham University Hospitals Trust, UK Sept 2016 to Sept 2017	• HM JACKarc + HM JACKarc analyser 2WW investigations	2WW patients who returned 2 types of FIT test	38/732 (5.19%)	4, 10, 22.6, 150	10	89.00 (75-97)	74.00 (70-77)	Using OC-S results in higher referrals. Consequently, OC-S detected more cancers than HM-J for the same cut-offs. Suggest that analyser-specific f-Hb cut-offs are needed, especially at lower f-Hb.
		• OC-Sensor DIANA					84.00 (69-94)	78.00 (75-81)	
2	Benton 2022[46] 50 NHS hospitals across England, UK Oct 2017 to Dec 2019	• HM-JACKarc colonoscopy	NG12 high risk, who had colonoscopy. Randomised to cohort 1 who were given 4 tests	7/233 (3.00%)	LoD, 10, 100	10	57.10 (25.1–84.2)	84.50 (79.2–88.6)	At 10 µg/g, < half the number of referrals would be made using SENTiFIT 270/FOB Gold Wide system compared to the other methods and dramatically fewer at the LoD. The calibration for the SENTiFIT 270/FOB Gold Wide gives lower f-
		• OC-Sensor PLEDIA				10	71.40 (35.9–91.8)	85.80 (80.7–89.8)	

	NICE FIT	• FOB Gold Wide - SENTiFIT 270				10	57.10 (25.1– 84.2)	93.40 (89.3– 95.9)	Hb results than the other three systems. Supported by Bland Altman Difference plot. Further work is required to understand the clinical impact of these differences and to minimise them.
		• NS-Prime				10	71.40 (35.9– 91.8)	83.60 (78.2– 87.9)	
3	MacLean 2022a[62] Royal Surrey Foundation Trust, UK July 2019 and March 2020	• FOB Gold Wide SENTiFIT 270 Colonoscopy or CTC or flexisig	2WW NG12 high/medium-risk	14/553 (2.53%)	10, 100, 150	10	100.00 (78.5 – 100)	84.80 (81.5 - 87.6)	Good agreement around negative threshold, but more patients would be triaged to further colonic investigation if using the QuikRead go®
		• QuikRead go					92.90 (68.5 - 98.7)	70.10 (66.1 - 73.8)	

2WW, two week wait; 95% CI, 95% confidence interval; CRC, colorectal cancer; LoD, limit of detection; N, number; NG12, national guideline 12; OC-S, OC-Sensor

4.3.12 Subgroup analyses by patient characteristics

Exploration of the potential reasons for heterogeneity in diagnostic test accuracy across studies using meta-regression was considered. However, study level covariates relating to patient characteristics of interest were not reported in all studies. Instead, studies which reported diagnostic test accuracy for subgroups of patients were considered in subgroup analyses (see Sections 4.3.12.1 to 4.3.12.6).

4.3.12.1 Anaemia

Studies reporting data on anaemia are summarised in Table 21. Considering all the available data on anaemia regardless of the test used, population type and reference standard, eleven studies[55, 57, 69-71, 74, 75, 78, 79, 81, 85] reported data on anaemia or IDA. The studies can be broadly categorised as comparative, comparing those with anaemia to those without; comparative, comparing those with anaemia to the study population unselected on the basis of anaemia (whole cohort); or non-comparative.

When considering studies that compare those with anaemia to those without, both[55, 74] reported lower sensitivity and specificity at a threshold of 10 for those with anaemia. One further study[85] reported that the optimal threshold (defined as the point on the ROC curve that maximises sensitivity and specificity) for those with anaemia is higher than those without. It should be noted that the definition of "optimal" in this study is not necessarily the same as optimising the threshold for cost-effectiveness or clinical decision making, where it may be preferable to optimise either sensitivity (where the test is used to rule out disease) or specificity (where the tests is used to rule in disease).

Amongst studies that compared those with anaemia to the whole study population, the results were more mixed. One study[78] showed the same trend of lower sensitivity and specificity, three[69, 70, 75] showed higher sensitivity and lower specificity, and one[57] showed lower sensitivity and higher specificity.

Of particular note is a study by Withrow *et al*. 2022[70], which shows that sensitivity increases as the threshold for anaemia is increased (i.e., more anaemic), whilst the specificity decreases in both men and women.

It should be noted that the definition of anaemia varied across studies, and some considered IDA anaemia whilst other considered other types of anaemia as well or instead.

Table 21: Sensitivity and specificity of studies reporting data for patients with anaemia

Author, year Test	Population type	Anaemia type	Threshold, µg/g	N with CRC/N analysed	Sensitivity (95% CI)	Specificity (95% CI)	Summary of IDA/anaemia vs comparator
Studies comparing those with to those without anaemia/IDA							
Cunin 2020[74a] HM-JACKarc	2	**Whole cohort**	10	48/928 (5.2%)	85.4 (NR)	86.9 (NR)	**Vs "no IDA"** Sens & Spec lower
		No IDA	10	28/739	89.00 (70-97.1)	84.00 (81.1-86.6)	
		IDA	10	20/189	80.00 (55.7-93.3)	81.60 (74.8-87)	
Johnstone 2022a[55] HM-JACKarc	**1**	**No anaemia**	10	32/3238 (0.99%)	96.9 (96.3,97.5)[a]	81.3 (80,82.6)[a]	Sens & spec lower
		Anaemia	10	26/793 (3.28%)	84.6 (82.1,87.1)[a]	72.9 (69.8,76)[a]	
Turvill 2021[85] HM-JACKarc	**4**	**No IDA**	19	101/3582 (2.8%)	88.1 (80.2-93.7)	85.3 (84.0-86.4)	Optimal threshold higher Optimal FIT threshold was ≥21 vs ≥19µg/g in anaemic vs non-anaemic
		IDA	21	34/559 (6.1%)	82.40 (65.5-93.2)	81.50 (77.9-84.8)	
Turvill 2021[85] HM-JACKarc	**4**	**No non-ID anaemia**	19	110/3597 (3.1%)	84.5 (76.4-90.7)	85.0 (83.7-86.1)	Optimal threshold higher Optimal FIT threshold was ≥30 vs ≥19µg/g in anaemic vs non-anaemic
		Non-ID Anaemia	30	25/544 (4.6%)	92.00 (74.0-99.0)	85.50 (82.2-88.5)	
Studies comparing those with anaemia/IDA to patients unselected on the basis of anaemia (whole cohort)							
D'Souza 2021a[75] HM-JACKarc	**4**	**Whole cohort**	10	12/298 (4.03%)	92.20% (88.2, 95.2)	82.30% (81.3, 83.2)	Sens higher, spec similar
		IDA	10	16/479 (3.34%)	100% (89.4, 100)	81.60% (77.7, 85.1)	
Tang 2022[69] HM-JACKarc	**4**	**Whole cohort**	10	20/603 (3.32%)	90.00 (68.3–98.77)	83.20 (79.9-86.14)	Sens higher, spec lower (low events in IDA)

		IDA	10	1/78 (1.28%)	100 (NE,NE) [a]	76.6 (67.2,86) [a]	
Juul 2018[57] OC-Sensor	**4**	**Whole cohort**	10	54/3462 (1.56%)	94.4 (93.6,95.2) [a]	85.7 (84.5,86.9) [a]	Sens lower, spec higher/similar
		Unexplained anaemia	10	54/3462 (1.56%)	20.4 (16.6,24.2) [a]	79.5 (75.7,83.3) [a]	
Gerrard 2023[78] Single FIT, HM-JACKarc	**1**	**Whole cohort**	10	69/2260 (3.1%	84.10 (73.3-91.8)	77.4 (75.6-79.1)	Sens similar/lower, spec lower
		Anaemia	10	38/567 (6.70%)	81.6 (78.4,84.8) [a]	68.6 (64.8,72.4) [a]	
Gerrard 2023[78] Dual FIT, HM-JACKarc	**1**	**Whole cohort**	10	88/2637 (3.3%)	96.60 (90.4-99.3) [a]	71.2 (69.4-73.0) [a]	Sens similar/lower, spec lower
		Anaemia	10	29/480 (6.04%)	93.1 (90.8,95.4) [a]	60.1 (55.7,64.5) [a]	
Withrow 2022[70] HM-JACKarc	**4**	**Whole cohort (both sexes)**	10	139/16604 (0.84%)	92.1 (91.7,92.5) [a]	91.5 (91.1,91.9) [a]	Sens higher, spec lower
		Low Haemoglobin (<130 g/L in men, <120g/L in women)	10	72/507 (1.42%)	95.8 (95.2,96.4) [a]	88 (87.1,88.9) [a]	
		Whole cohort (men)	10	83/7019 (1.18%)	92.8 (92.2,93.4) [a]	90.3 (89.6,91) [a]	Sens same or higher, spec lower with increasing anaemia
		Men, <130 g/L	10	46/2091 (2.20%)	93.5 (92.4,94.6) [a]	85.5 (84,87) [a]	
		Men, <120 g/L	10	36/1141 (3.16%)	91.7 (90.1,93.3) [a]	82.7 (80.5,84.9) [a]	
		Men, <110 g/L	10	23/494 (4.66%)	95.7 (93.9,97.5) [a]	79 (75.4,82.6) [a]	
		Men, <100 g/L	10	14/216 (6.48%)	100 (NE,NE) [a]	72.3 (66.3,78.3) [a]	
		Men, <90 g/L	10	9/89 (10.11%)	100 (NE,NE) [a]	71.2 (61.8,80.6) [a]	
		Whole cohort (women)	10	57/9585 (0.59%)	91.1 (90.5,91.7) [a]	92.4 (91.9,92.9) [a]	Sens higher, spec lower with increasing anaemia

		Women, <120g/L	10	25/2758 (0.91%)	100 (NE,NE) [a]	89.4 (88.3,90.5) [a]	
		Women, <110g/L	10	13/1297 (1.00%)	100 (NE,NE) [a]	88 (86.2,89.8) [a]	
		Women, <100g/L	10	6/491 (1.22%)	100 (NE,NE) [a]	84.5 (81.3,87.7) [a]	
		Women, <90g/L	10	3/189 (1.59%)	100 (NE,NE) [a]	79.6 (73.9,85.3) [a]	
Non-comparative studies							
Ayling 2019[71] OC-Sensor	4	Anaemia	10	7/178 (3.93%)	71.4 (64.8,78) [a]	95.9 (93,98.8) [a]	
Ayling 2019[71] OC-Sensor	4	IDA	10	6/137 (4.38%)	68.7 (60.9,76.5) [a]	95.4 (91.9,98.9) [a]	
Morales-Arraez 2018[79] OC-Sensor	4	Moderate-severe IDA	10	28/245 (11.43%)	92.9 (89.7,96.1) [a]	57.1 (50.9,63.3) [a]	
Rodriguez-Alonso 2020[81]	4	IDA	10	9/120 (7.5%	100 (NE,NE) [a]	77.5 (70,85) [a]	

95% CI, 95% confidence interval; CRC, colorectal cancer; IDA, iron deficiency anaemia; N, number; NE, not estimable
[a]calculated by EAG reviewer

4.3.12.2 Age

Three studies[70, 76, 85] reported data according to age groups (see Table 22). All were large studies with >5000 patients, the largest included 16,604 patients.[70] All studies used HM JACKarc.

One study[85] reported the optimal cut off (the point that maximises both sensitivity and specificity) based on the ROC curves for those aged under 60 years and aged 60+ years separately, and reported that the optimal threshold was lower in the 60+ age group (19 µg/g) compared to the younger age group (37 µg/g). This study concluded that FIT could be incorporated into a risk score based on sex, age, symptoms and signs, drug history, and blood parameters, but did not conduct the analyses required to produce such a score in that publication. It should be noted that the definition of "optimal" in this study is not necessarily the same as optimising the threshold for cost-effectiveness or clinical decision making, where it may be preferable to optimise either sensitivity (where the test is used to rule out disease) or specificity (where the tests is used to rule in disease).

Another study[76] reported a limited range of thresholds (2, 10 and 150 µg/g) for those aged under 50 years and aged 50+years. Sensitivity was lower in the younger age group at any given threshold though confidence intervals overlapped. This trend was less evident at the highest threshold and the number of events in the younger age group was small (n=16). In this study the authors noted that in younger patients it may be appropriate to intepret any detectable faecal Hb as a positive test.

In the third study,[70] the thresholds were 2 µg/g and 10 µg/g and were reported for those aged under 40 years, then for those aged 40+, 50+, 60+, 70+ and 80+ years. This study performed multivariable modelling including FIT, blood tests, age, and sex and concluded that that age-specific thresholds for FIT positivity would not improve test performance.

Overall, there is some indication that FIT thresholds may need to be lower in younger patients in order to achieve the same sensitivity as for older patients. The available data does not provide conclusive evidence that different FIT thresholds should be used or what they should be.

Table 22: Sensitivity and specificity by age

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Popul-ation type	N with CRC/ N analysed (%)	Thres-hold, µg/g	Age group in years	Sensitivity (95% CI)	Specificity (95% CI)	Conclusion drawn by study authors
1	D'Souza 2021b[76] NICE FIT October 2017 to December 2019	HM JACKarc analytical system Colonoscopy	**4**	16/1103 (1.45%)	2	<50	87.50 (61.7–98.4)	70.40 (67.6–73.1)	Detectable f-Hb on FIT in symptomatic younger patients may indicate referral for investigation of colorectal cancer and serious bowel disease.
				313/8719 (3.59%)	2	50+	97.40 (95.0–98.9)	64.10 (63.1–65.2)	
				16/1103 (1.45%)	10	<50	81.30 (54.4–96.0)	83.60 (81.3–85.5)	
				313/8719 (3.59%)	10	50+	91.40 (87.7–94.2)	83.50 (82.7–84.3)	
				16/1103 (1.45%)	150	<50	68.80 (41.3–89.0)	92.20 (90.4–93.7)	
				313/8719 (3.59%)	150	50+	70.90 (65.6–75.9)	94.90 (94.4–95.3)	
2	Turvill 2021[85] Yorkshire & Humber, UK April 2018 to Dec 2019 Fast track FIT	HM JACKarc Full colonoscopy or CT colonography, or a lesser investigation (such as CT abdomen/pelvis with contrast or flexible sigmoidoscopy)	**4**	30/1217 (2.5%)	37	<60	90.00 (73.5-97.9)	87.40 (85.4-89.3)	The optimal cut-off value for people aged ≥60 years (19 µg/g faeces) is lower than for those aged <60 years (37 µg/g faeces). FIT could be incorporated into a risk score based on sex, age, symptoms and signs, drug history, and blood parameters
				19/3823 (0.49%)	19	60+	83.50 (75.6-89.6)	85.40 (84.2-86.5)	

3	Withrow 2022[70] (same study as Nicholson 2020)[29] Oxfordshire, UK March 2017 to December 21, 2020 CSS-BIO-3 4730	HM JACKarc Records follow-up	**4**	9/1390 (0.65%)	2	<40	100 (70.1 - 100)	89.1 (87.4 to 90.7)	The lack of an apparent age-effect after taking into account FIT suggests that age-specific thresholds for FIT positivity would not improve test performance
				130/15214 (0.85%)	2	>40	96.2 (91.3-98.3)	83.0 (82.4-83.6)	
				118/12936 (0.91%)	2	>50	95.8 (90.5-98.2)	81.8 (81.1-82.4)	
				98/8755 (1.12%)	2	>60	94.9 (88.6-97.8)	78.8 (77.9-79.7)	
				77/3043 (2.53%)	2	>70	94.8 (87.4-98)	51.8 (50-53.6)	
				41/2527 (1.62%)	2	>80	95.1 (83.9-98.7)	68.4 (66.6-70.2)	
				9/1390 (0.65%)	10	<40	88.9 (56.5-98)	93.4 (92-94.6)	
				130/15214 (0.85%)	10	>40	92.3 (86.4-95.8)	91.3 (90.9-91.8)	
				118/12936 (0.91%)	10	>50	91.5 (85.1-95.3)	90.7 (90.2-91.2)	
				98/8755 (1.12%)	10	>60	89.8 (82.2-94.4)	89.0 (88.3-89.6)	
				77/5863 (1.31%)	10	>70	89.6 (80.8-94.6)	87.1 (86.2-87.9)	
				41/2533 (1.62%)	10	>80	87.8 (74.5-94.7)	83.1 (81.6-84.5)	

95% CI, 95% confidence interval; CRC, colorectal cancer; N, number

4.3.12.3 Sex

Three studies[29, 45, 85] reported data for men and women separately (see Table 23). All were studies with >3000 patients, with the largest including 9,899 patients.[29] One study used OC-Sensor PLEDIA[45] and two used HM-JACKarc.[29, 85]

One study[85] reported the optimal cut off (the point that maximises both sensitivity and specificity), based on the ROC curves for men and women separately, and reported that the optimal threshold was lower for women (16 µg/g) than for men (21 µg/g). This study concluded that FIT could be incorporated into a risk score based on sex, age, symptoms and signs, drug history, and blood parameters. It should be noted that the definition of "optimal" in this study is not necessarily the same as optimising the threshold for cost-effectiveness or clinical decision making, where it may be preferable to optimise either sensitivity (where the test is used to rule out disease) or specificity (where the tests is used to rule in disease).

The two other studies[45, 70] reported a range of thresholds (from 10 to 150µg/g), and generally showed that at thresholds above 10 µg/g, sensitivity and specificity is higher in women than in men. This difference was more pronounced in one study[45] than the other,[29] but due to the small number of studies it was not possible to tell if this was due to the use of different analysers or some other factor. At 10µg/g, one study showed roughly equivalent sensitivity and specificity,[29] whilst the other study showed numerically lower sensitivity in men, but stated that no significant difference in FIT sensitivity was found.[45] Withrow 2022 conducted a multivariable analysis including sex and showed the probability of colorectal cancer reached 3% at 17 and 25 µg/g for males and females, respectively.

If sensitivity and specificity are different in women than in men at a given threshold, a different threshold in women may be required to achieve equivalent sensitivity and specificity in the two sexes. However, it was not possible on the basis of the available data to conclude what and whether different FIT cut-off values are required according to sex.

Table 23: Sensitivity and specificity by sex

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	N with CRC/ N analysed (%)	Threshold, µg/g	Men: Sensitivity (95% CI)	Men: Specificity (95% CI)	Women: Sensitivity (95% CI)	Women: Specificity (95% CI)	Conclusion drawn by study authors
1	**Ball 2022[45] (additional data by personal communication)** Sheffield, UK Oct 2019 to Dec 2019	OC-Sensor PLEDIA Colonoscopy or CT imaging and colon capsule endoscopy	4	Men: 25/1566 (1.6%) Women: 20/1940 (1.03%)	10	84.00 (63.1–94.7)	79.20 (77.0–81.2)	100.00 (80 -100)	82.00 (80.2–83.7)	Sex did not significantly influence FIT sensitivity on subgroup analysis.
					20	80.00 (58.7–92.4)	85.40 (83.5–87.1)	95.00 (73.1–99.7)	88.80 (87.2–90.2)	
					50	68.00 (46.4–84.3)	91.60 (90.0–92.9)	80.00 (55.7–93.3)	94.10 (92.9–95.1)	
					80	64.00 (42.6–81.3)	93.90 (92.6–95.0)	70.00 (45.7–87.2)	95.80 (94.8–96.6)	
					100	64.00 (42.6–81.3)	94.60 (93.3–95.7)	70.00 (45.7–87.2)	96.70 (95.8–97.4)	
					120	60.00 (38.9–78.2)	95.20 (94.0–96.2)	65.00 (40.9–83.7)	97.00 (96.1–97.7)	
					150	52.00 (31.8–71.7)	96.40 (95.3–97.2)	55.00 (32.0–76.2)	97.30 (96.5–98.0)	
2	Turvill 2021[85] Yorkshire & Humber, UK	HM JACKarc Full colonoscopy or CT colonography,	**4**	Men: 89/2242 (3.9%)	Men: 21 Women: 16 NB: optimal	85.40 (76.3 to 92.0)	83.70 (82.0 to 85.2)	87.10 (76.1 to 94.3)	85.60 (84.2 to 86.9)	The optimal cut-off value is lower for females (16 µg/g faeces) than for males (21 µg/g faeces. FIT

	April 2018 to Dec 2019 Fast track FIT	or a lesser investigation (such as CT abdomen/pelvis with contrast or flexible sigmoidoscopy)		Women: 62/2798 (2.2%)	threshold was derived					could be incorporated into a risk score based on sex, age, symptoms and signs, drug history, and blood parameters
3	Nicholson 2020 (same study as Withrow 2022)[29a] Oxfordshire, UK March 2017 to December 21, 2020 CSS-BIO-3 4730	HM JACKarc Records follow-up	**4**	Men: 65/4104 (1.6%) Women: 40/5795 (0.69%)	7	92.30 (85.8-98.8)	87.90 (86.9-88.9)	90.00 (80.7-99.3)	91.10 (90.3-91.8)	The area under the curve for all adults did not change substantially by gender. From Withrow 2022: The probability of colorectal cancer reached 3% at 17 and 25 µg/g, for males and females respectively.
					10	90.80 (83.7-97.8)	89.80 (88.8-90.7)	90.00 (80.7-99.3)	92.40 (91.8-93.1)	
					20	83.10 (74.0-92.2)	92.30 (91.5-93.2)	87.50 (77.3-97.7)	94.60 (94.1-95.2)	
					50	73.80 (63.2-84.5)	95.50 (94.9-96.2)	75.00 (61.6-88.4)	96.90 (96.5-97.4)	
					100	60.00 (48.1-71.9)	96.80 (96.3-97.3)	62.50 (47.5-77.5)	98.10 (97.8-98.5)	
					120	55.40 (43.3-67.5)	97.20 (96.7-97.7)	60.00 (44.8-75.2)	98.30 (98.0-98.6)	
					150	50.80 (38.6-62.9)	97.50 (97.1-98.0)	60.00 (44.8-75.2)	98.50 (98.2-98.8)	

95% CI, 95% confidence interval; CRC, colorectal cancer; N, number

[a] Nicholson 2020 (n=9896) is an earlier data cut of the same study as Withrow 2022 (n=11,142). The data from Nicholson 2020 has been included in this analysis over the Withrow 2022 data as it reports more thresholds, even though the study population is smaller. However, the Withrow 2022 study conducted a multivariable analysis including sex, and the conclusions relating to this have been reported.

4.3.12.4 Medications that might cause GI bleeding

The scope issued by NICE states the assessment should consider whether the FIT threshold should be different for "People taking medications or with conditions which increase the risk of gastrointestinal bleeding". In a slight widening of the scope, NICE confirmed an additional study which looked at the effect of taking proton pump inhibitors (PPI), which may decrease the risk of gastrointestinal bleeding, may be of interest to the committee and it has therefore been included.

Consequently, three studies[73, 80, 85] were included in this subgroup analysis. The studies are summarised in Table 24. All studies included more than 1000 patients, with the largest including 5040 in total.[85] Two studies (three references) used OC-Sensor analysers,[73, 80, 86] and one used HM-JACKarc.[85]

One study[85] reported the optimal cut off (the point that maximises both sensitivity and specificity) based on the ROC curves for those using antiplatelet, anticoagulants NSAIDs and those not using these drugs. The optimal threshold was 19µg/g in both cases, though the sensitivity and specificity were superior in those not using the drugs than in those who were. This study concluded that FIT could be incorporated into a risk score based on sex, age, symptoms and signs, drug history, and blood parameters.

Another study (part of the "colonpredict" study)[73, 86] reported test accuracy data for those using aspirin and those not using aspirin. It should be noted that this study recruited symptomatic patients from secondary as well as primary care and was therefore excluded from the main analysis. The analysis of aspirin users was included due to the sparsity of data in this subgroup, but it is unclear how generalisable these results will be to the primary care setting. Only one threshold was included (20 µg/g). As with the previous study, the sensitivity and specificity were superior in those not using the drug. This study concluded that aspirin use did not change the diagnostic accuracy of FIT in patients with gastrointestinal symptoms.

The third study compared PPI users to PPI non-users. It should be noted that this study recruited symptomatic patients from secondary as well as primary care and was therefore excluded from the main analysis. The analysis of PPI users was included due to the sparsity of data in this subgroup, but it is unclear how generalisable these results will be to the primary care setting. At a threshold of 20 µg/g sensitivity was similar, and specificity was slightly higher in non-users. This study did not conclude anything for the detection of CRC, but concluded there was impaired FIT performance in PPI users for the detection of advanced neoplasia.

Conclusion: The evidence base is currently small, and it was not possible on the basis of the available data to conclude what and whether different FIT cut-off values are required according to medications being taken by a patient.

Table 24: Sensitivity and specificity for patients taking medications that may affect the risk of GI bleeding

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Inclusion criteria	Group	N with CRC/ N analysed (%)	Thres-hold, µg/g	Sensitivity (95%CI)	Specificity (95%CI)	Conclusion drawn by study authors
1	Bujanda 2018[73] Spain (assume Ourense and San Sebastian)[86] March 2012 to 2014 COLONPREDICT	• OC-Sensor[86] Colonoscopy	Population type 4 - Symptomatic patients referred from primary and secondary care	Aspirin users	51/485 (10.51%)	20	88.00 (75-95)	66.97 (62-71)	Aspirin use did not change the diagnostic accuracy of FIT in patients with gastrointestinal symptoms.
				Aspirin non-users	299/2567 (11.65%)	20	92.00 (88-95)	71.00 (69-73)	
2	Turvill 2021[85] Yorkshire & Humber, UK April 2018 to Dec 2019 Fast track FIT	HM JACKarc Full colonoscopy or CT colonography, or a lesser investigation (such as CT abdomen/pelvis with contrast or flexible sigmoidoscopy)	Population type 4 - 2WW patients	Antiplatelets, anticoagulants NSAIDs	19/1356 (1.4%)	19	82.40 (69.1-91.6)	80.50 (78.2-82.6)	The specificity differed according to use of antiplatelets, anticoagulants NSAIDs. FIT could be incorporated into a risk score based on sex, age, symptoms and signs, drug history, and
				No use of antiplatelets, anticoagulants NSAIDs	100/3684 (2.7%)	19	87.0 (78.8-92.9)	86.9 (85.7-88.0)	

									blood parameters
3	Rodriguez-Alonso 2018[80] Barcelona, Spain Sept 2011 to Oct 2012	OC-Sensor MICRO Colonoscopy	Population type 4 - Symptomatic patients referred from primary and secondary care	PPI users	15/525 (2.86%)	20	93.3	85.1	No conclusion drawn for the identification of CRC in PPI users, concluded impaired FIT performance in PPI users for detection of advanced neoplasia.
				PPI non-users	15/477 (3.14%)	20	93.3	87.4	

95% CI, 95% confidence interval; CRC, colorectal cancer; N, number

4.3.12.5 Ethnicity

No studies reporting the diagnostic test accuracy of any of the in-scope tests according to ethnicity were identified.

4.3.12.6 People with blood disorders

No studies reporting the diagnostic test accuracy of any of the in-scope tests in a subgroup of people with blood disorders (e.g., beta thalassemia) that could affect the performance of the test were identified.

4.3.13 Advanced adenomas and inflammatory bowel disease outcomes

Nine studies (10 publications)[17, 49, 57, 59, 61, 63-65, 78, 82] reported data on AA and IBD. These are summarised in Table 25. One study reported data for the IDK Hb and Hb/Hp complex ELISA tests, whilst the remainder reported data for immunoturbidimetry tests. The synthesis focussed on the immunoturbidimetry tests, but the IDK data was used in the model for the IDK tests.

4.3.13.1 Statistical synthesis of AA outcomes

Nine studies[17, 49, 57, 59, 61, 63-65, 78] contributed to the meta-analysis for AA outcomes (HM-JACKarc: 6, OC-Sensor: 2 QuikRead Go: 1). Five studies provided diagnostic accuracy at a single threshold and the maximum number of thresholds considered by a single study was 3. The full dataset (all studies) provided a total of 15 pairs of sensitivity and specificity, at thresholds between 2 and 150. Figure ***11***A and Figure ***11***B illustrate the results for all studies, irrespective of test type. Separate syntheses are also provided for HM-JACKarc (Figure ***11***C) and OC-Sensor (Figure ***11***D).

One of the studies[78] also reported data for AA and IBD when using Dual FIT and is reported in section 4.3.7.

For the analysis of all test types together, sensitivity ranges from 80.4 (95% CrI: 55.8, 98.3; 95% PrI: 50.0, 100.0) at a threshold of 2, to 20.4 (95% CrI: 0.6, 47.5; 95% PrI: 0, 57.4) at a threshold of 150. Specificity ranges from 51.6 (95% CrI: 31.6, 71.1; 95% PrI: 3.2,98) at a threshold of 2, to 95.7 (95% CrI: 82.5, 99.5; 95% PrI: 58.5,100) at a threshold of 150. There is a large amount of heterogeneity between studies, as illustrated by the wide 95% CrI and PrI. Point estimates of summary sensitivity and specificity changed considerably for the separate analyses by test type (see Figure 11 C, D and Table 26), emphasising the large amount of uncertainty.

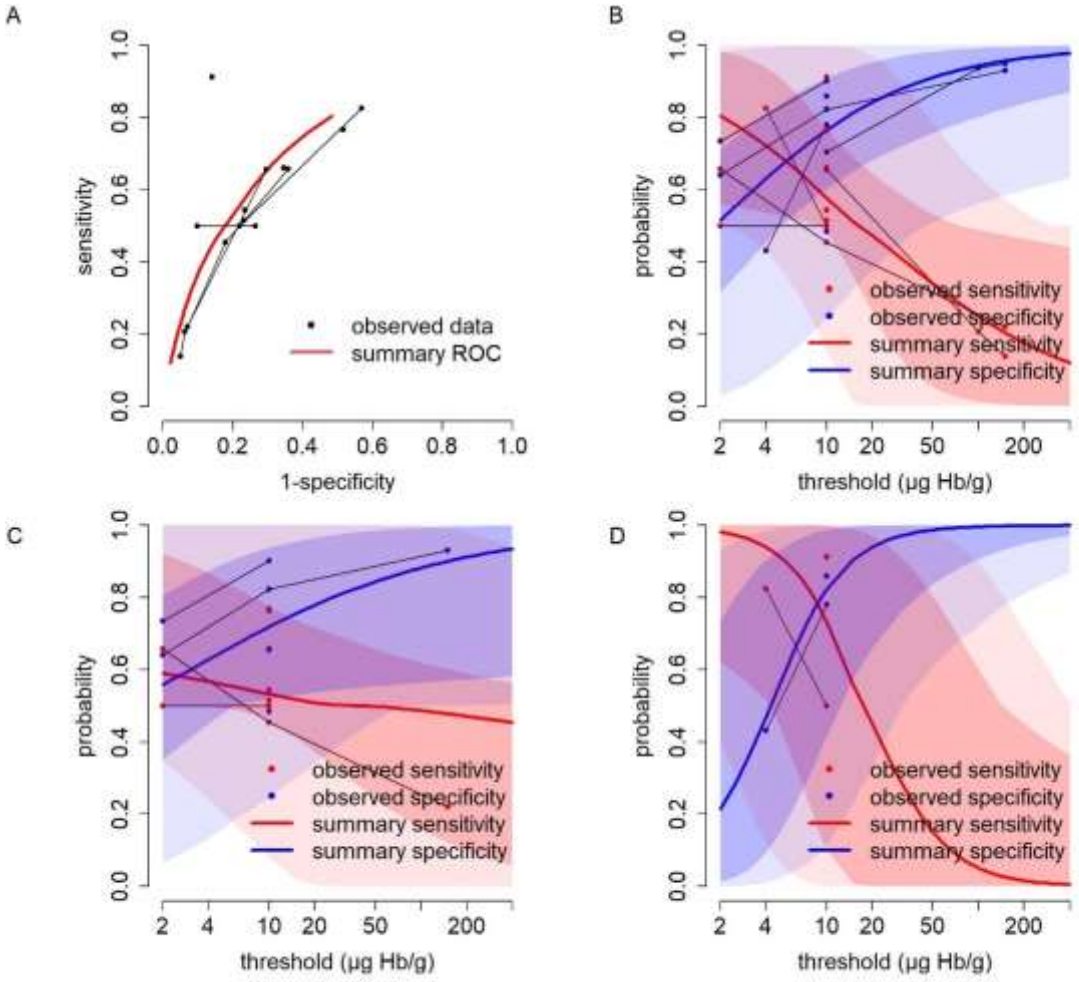
Table 25: Studies reporting data on AA and IBD

#	Author, year Location Recruitment dates Study name (if available)	Analyser Reference standard	Population types	N with AA / N analysed (%) N with IBD / N analysed (%)	Thresholds, µg/g
1	Sieg 1999[82] Ostringen, Germany NR, prior to publication in 1999	Immunological test for Hb/Hp complex Colonoscopy Immunological test for HB Colonoscopy	4	AA: 37/621 (5.95%) IBD: 22/621 (3.5%)	2
2	D'Souza 2020a[49] Croyden, UK Nov 2016 to Oct 2017	HM JACKarc analytical system Colonoscopy	1,2,3	AA (population 1): 4/298 (1.3%) IBD (population 1): 12/298 (4.0%)	2, 10
3	D'Souza 2021a[75] D'Souza 2021c[17d] NICE FIT October 2017 to December 2019	HM JACKarc analytical system Colonoscopy	4	AA: 421/982 (4.3%) IBD: 427/9822 (4.3%)	2, 10, 150

4	Gerrard 2023[78] Lothian, Scotland, UK Jan 2019 to Feb 2020	HM-JACKarc Endoscopy or CT with colorectal protocol.	1	AA: 105/2260 (4.6%) and 136/3426 (4.0%) IBD: 59/226 (2.6%) and 55/3426 (1.6%)	10
5	Juul 2018[57] Central Denmark Sept 2015 to Aug 2016 NCT02308384	OC-Sensor DIANA Records follow-up	4	AA: 68/3462 (1.9%) IBD: 31/3462 (0.9%)	10
6	MacDonald 2022[59] NHS Lanarkshire, Scotland, UK October 2016 to February 2019	HM-JACKarc Records follow-up	1	AA: 47/5250 (0.9%) IBD: 131/5250 (2.5%)	10
7	MacLean 2021b Royal Surrey Foundation Trust, UK July 2019 and March 2020	QuikRead go Colonoscopy, CTC or flexisig	2	AA: 29/553 (5.2%) IBD: 9/553 (1.6%)	10, 100, 150
8	Mowat 2016[65] NHS Tayside, Scotland, UK Oct 2013 to March 2014	OC-Sensor iO Colonoscopy	4	AA: 40/750 (5.3%) IBD: 34/750 (4.5%)	4, 10
9	Mowat 2021[64] & 2019[63] NHS Tayside, Scotland, UK December 2015 to December 2016	HM JACKarc Records follow-up	4	AA: 133/1447 (9.2%) IBD: 68/1447 (4.7%)	10

AA, advanced adenomas; IBD, inflammatory bowel disease; N, number

Figure 11: Observed data and summary sensitivity and specificity for AA outcomes. A) All tests, ROC B) All tests as a function of threshold, C) HM-JACKarc D) OC-Sensor



Results for OC-Sensor use an informative prior, based on the synthesis of all tests together.

Table 26: Summary sensitivity and specificity at selected thresholds for AA outcome

threshold, µg/g	All tests (S=9)		HM-JACKarc (S=6)		OC-Sensor (S=2)	
	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity
2	80.4 (55.8,98.3)	51.6 (31.6,71.1)	59.1 (50,92)	55.9 (35.1,80.6)		
2.5	78 (55.2,97.4)	55.4 (36.1,74.4)	58.4 (50,90.5)	58.1 (38.4,82.9)		
3	75.9 (54.7,96.4)	58.4 (39.6,77.1)	57.7 (50,89.3)	59.9 (41,84.6)		
4	72.2 (53.7,93.8)	63.1 (45.6,81)	56.7 (49.8,86.8)	62.8 (45.1,87.3)	93.9 (51.5,100)	46.8 (9.5,90.3)
7	63.9 (51.4,84.6)	71.7 (55.3,87.7)	54.7 (48.2,81.2)	68.5 (50.8,91.7)	84.6 (27.8,100)	70.7 (27.2,96.7)
10	57.7 (48.6,76.7)	76.5 (60.3,90.9)	53.2 (45.9,77.6)	71.9 (52,93.8)	73.2 (10.1,99.9)	82.2 (41.6,98.7)
20	47.4 (26.1,64.4)	84.2 (68.1,95.3)	50.9 (37.3,71.6)	77.9 (53.7,96.5)		
50	34.1 (5.6,53.2)	91.1 (75.7,98.2)	49.8 (24.3,65.7)	84.4 (55.4,98.5)		
100	25 (1.4,48.9)	94.4 (80.2,99.2)	48.7 (16,61.9)	88.2 (56.3,99.2)		
120	22.8 (1,48.3)	95 (81.3,99.3)	48.3 (14.2,61.1)	89.1 (56.6,99.4)		
150	20.4 (0.6,47.5)	95.7 (82.5,99.5)	47.8 (12.3,60.1)	90.1 (56.9,99.5)		

4.3.13.2 Statistical synthesis of Inflammatory bowel disease outcomes

Nine studies contributed to the meta-analysis for IBD outcome (HM-JACKarc: 6, OC-Sensor: 2 QuikRead go: 1). 5 provided diagnostic accuracy at a single threshold and the maximum number of thresholds considered by a single study was 3. The full dataset (all studies) provided a total of 15 pairs of sensitivity and specificity, at thresholds between 2 and 150.

Figure 12A and B illustrates the results for all studies, irrespective of test type. Separate syntheses are also provided for HM-JACKarc (Figure 12 C) and OC-Sensor (Figure 12 D).

For the analysis of all test types together, sensitivity ranges from 85.7 (95% CrI: 70, 96.7; 95% PrI: 42.3, 100.0) at a threshold of 2, to 41.7 (95% CrI: 15.9, 66.1; 95% PrI: 0.9, 91.4) at a threshold of 150. Specificity ranges from 53.8 (95% CrI: 33.1, 75.5; 95% PrI: 2.6, 99.3) at a threshold of 2, to 95.0 (95% CrI: 80.2, 99.5; 95% PrI: 55.0,100) at a threshold of 150. As with AA, there is a large amount of heterogeneity between studies, as illustrated by the wide 95% CrI and PrI. Point estimates of summary sensitivity and specificity changed considerably for the separate analyses by test type (see Figure 12 C, D and Table 27), emphasising the large amount of uncertainty.

Figure 12: Observed data and summary sensitivity and specificity for All tests. IBD outcomes

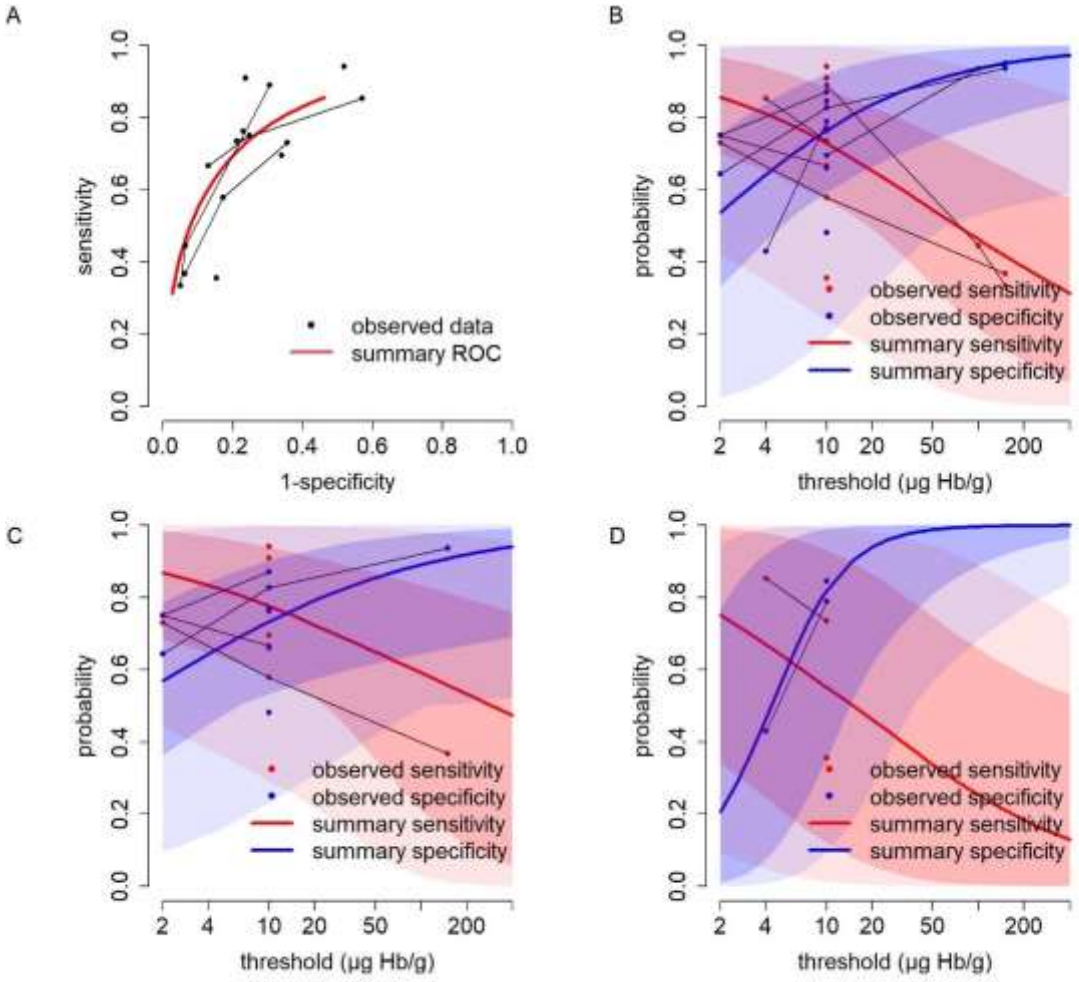


Table 27: Summary sensitivity and specificity at selected thresholds for IBD outcomes

threshold, µg/g	All tests (S=9)		HM-JACKarc (S=6)		OC-Sensor (S=2)	
	sensitivity	specificity	sensitivity	specificity	sensitivity	specificity
2	85.7 (70,96.7)	53.8 (33.1,75.5)	86.8 (68.4,98.5)	57 (36.4,78.7)		
2.5	84.3 (68.5,96)	57.2 (37.2,78.4)	85.8 (67.3,98.1)	59.4 (39.2,80.6)		
3	83.1 (67.2,95.3)	60 (40.5,80.6)	84.9 (66.4,97.8)	61.3 (41.3,82.1)		
4	81 (65.1,94)	64.2 (45.8,84)	83.4 (64.7,97.1)	64.3 (44.8,84.4)	67 (24.7,97.9)	46.4 (7.4,92)
7	76.3 (60.4,90.7)	72 (54.7,89.4)	80.1 (61.2,95.3)	69.9 (50.8,88.1)	59.8 (16.4,95.5)	70.3 (22.3,97.5)
10	72.9 (57.1,88.2)	76.4 (59.2,92.1)	77.6 (58.6,94)	73.3 (53.3,90.2)	55.1 (12.2,93.1)	81.9 (35.3,99)
20	65.3 (49.2,82.9)	83.6 (66.3,95.8)	72.3 (52.2,91.1)	79.2 (57.3,93.4)		
50	54.4 (33.6,75.5)	90.3 (73.7,98.3)	64.9 (35.1,86.9)	85.4 (61.7,96.2)		
100	46.3 (21.7,69.7)	93.6 (78,99.2)	59.2 (21.3,83.4)	89.1 (64.3,97.5)		
120	44.2 (19,68.1)	94.3 (79,99.3)	57.7 (18.3,82.4)	89.9 (64.9,97.8)		
150	41.7 (15.9,66.1)	95 (80.2,99.5)	55.7 (15.2,81.2)	90.8 (65.8,98.1)		

4.3.14 other outcomes

4.3.14.1 Test uptake and repeat tests

Test failures, uptake and repeat tests

Since these outcomes are likely to be affected by the point within the care pathway at which FIT is issued to the patient, this analysis has been restricted to studies where FIT was issued in primary care. All Dual FIT studies were conducted in secondary care, and have been included as no other data was available. The data is summarised in Table 28. Additional data with lower generalisability (studies in secondary care settings and that asked patients to provide samples for multiple tests) are provided in Appendix 9.

Test failure rates: Test failure rates were reported by ten studies (11 references)[44, 45, 47, 52, 55, 57, 59, 63-65, 71] ranged from 0.2% in two[44, 65] separate studies, to 18.8%.[52] Data were available for OC-Sensor and HM-JACKarc only; there was no strong evidence that rates differed according to these test brands. The majority of studies reported rates between 2% and 5%.[47, 55, 57, 59, 63, 64, 71] The study with the highest value[52] was not obviously different in nature from other studies within the analysis where FIT was used in routine practice, e.g. Mowat *et al.* 2021. [63, 64] However, it was not clear if all studies defined this outcome consistently. The study that had the highest value also provided the most details about the test failures, which included problems such as labelling errors, incorrect containers, no date of collection,

volume errors and laboratory accidents.[52] Other studies tended to report spoiled or unsuitable samples, which may represent a narrower definition of test failure, though a precise definition was often missing.

One study[78] in Dual FIT reported that FIT was inappropriate for 4.5% of patients, or that emergency presentation predated FIT postage.

Uptake: Only two[44, 65] studies in primary care explicitly reported non-return of FIT, both in OC-Sensor. One had an extremely high non-return rate (52%),[65] but this may be confounded by the fact that a referral had already been made and did not depend on the return of the FIT sample. The other study reported non-return rate of 9.4%, where FIT was being used as part of the diagnostic pathway. A later update[72] of the same study reported 3631/38920 (9.3%) first FIT requests were not returned.

One study[78] in Dual FIT showed 10.7% returned no FIT, and a further 20.5% returned only one FIT. Another[54] reported 4.9% only returned one FIT, and one further study[83] noted stool sample was missing for 16.1% of patients. All studies took place in secondary care.

Repeat tests: Four studies (5 references)[44, 45, 56, 63, 64] reported data on repeat FIT tests. The largest of these was a study pooling data from three Scottish regions. Of 135,396 tests, 12,359 (9.1%) were repeat FITs. This study also reported how many times repeat FITs were ordered for some patients, as can be seen in column 7 of Table 28. The other three studies report that 0.7%,[45] 1.7%[44] and 2.07% [63, 64] repeat FITs were ordered, though a later update[72] of one study[44] reported 8349 (17.0%) requests were repeat tests in 6640 patients.

Table 28: Studies issuing FIT in primary care or issuing DUAL FIT, and reporting test failure rates, test uptake and number of repeat tests

Author, year	Analyser	FIT provided in	N with CRC/ N analysed (%)	Invalid/ test failure rates	Test uptake /non-return	Repeat tests
Johnstone 2022a[55]	HM-JACKarc (personal communication)	Primary care	61/4737 (1.29%)	231/4968 (4.6%)	NR	NR
MacDonald 2022[59]	HM-JACKarc	Primary care, those undergoing referral	151/5250 (2.88%)	Rejected for technical reasons 115 (2.1%)	NR	NR
Mowat 2021[64] & 2019[63]	HM JACKarc	Primary care	105/5381 (1.95%)	Unsuitable for analysis, n=152/5422 (2.8%)	NR	n=112/5422 (2.07%) repeat tests
Johnstone 2022b[56] Symptomatic patients who had 2 FITs between 1 week and 1 year apart	HM-JACKarc	Primary care	42/5761 individuals (0.73%)	NR	NR	12,359/ 135 396 (9.1%) repeat FITS in total, from 5761 individuals. FITs between 1 week and 1 year apart: 2 FITs: 5027 3 FITs: n=649 4 FITs: n=71 5 FITs: n=10 6 FITs: n=4
Bailey 2021a[44] Bailey *in press*[72]	OC-Sensor iO	Primary care	15589 FIT requests (CRC NR)	34/15589 (0.2%) spoiled or not suitable for analysis	Kit not returned 1393/14 788 (9.4%) Updated analysis: 3631/38920 (9.3%)[72]	229/13361 (1.7%) Updated analysis: 8349 (17.0%) requests were repeat tests in 6640 patients from 40817 patients[72]
Cama 2022[47]	OC-Sensor iO	Primary care	74/5341 (1.39%)	No result returned in 2% of samples (n=13,466)	NR	NR

Author, year	Analyser	FIT provided in	N with CRC/ N analysed (%)	Invalid/ test failure rates	Test uptake /non-return	Repeat tests
Georgiou Delisle 2022a[52]	OC-Sensor iO	Primary care	61/4187 (1.46%)	Could not be processed: 948/5050 (18.8%)[b]	NR	NR
Ball 2022[45]	OC-Sensor PLEDIA	Primary care	17/2892 (0.6%)	n=599/4219 (14.2%) due to insufficient clinical details, sample errors, insufficient ID/Labelling[c]	NR	n=29/4219 (0.7%)
Juul 2018[57]	OC Sensor DIANA	Primary care	54/3462 (1.56%)	Invalid FITs = 91/3745 (2.4%)	NR	NR
Mowat 2016[65]	OC Sensor iO	Primary care	28/750 (3.73%)	n=5/2789 (0.2%) spoiled/unsuitable samples	FIT not returned: 1130/2173[a] (52.0%)	
Subgroups						
Ayling 2019[71]	OC Sensor	Secondary care	Low Haemoglobin group: 7/178 (3.93%) IDA group: 6/137 (4.38%)	6/184 (3.3%) FIT unusable	NR	NR
DUAL FIT						
Gerrard 2023[78]	HM-JACKarc	Secondary care	88/2637 (3.34%)	Clinician considered FIT inappropriate, or emergency presentation predated FIT postage: 205/4559 (4.5%)	FIT not returned: 464/4354 (10.7%) Only one FIT returned: 891/4354 (20.5%)	NR
Hunt 2022[54]	OC-Sensor	Secondary care	317/28622 (1.11%)	NR	Only returned one FIT:	NR

Author, year	Analyser	FIT provided in	N with CRC/ N analysed (%)	Invalid/ test failure rates	Test uptake /non-return	Repeat tests
					1482/30104 (4.9%)	
Tsapournas 2020[83]	QuikRead go	Secondary care	13/242 (5.37%)	NR	Stool sample missing n=57/355 (16.1%)	NR

CRC, colorectal cancer; N, number

[a] NB in this study, patients had already been referred, so there was less incentive to return the FIT test if referral depended on FIT sample. Also had to do two tests on one sample (one FIT, one faecal calprotectin)

[b] Reason for incorrect FIT processing: Sample labelling errors, n=223 (5.3%); Wrong sample type, n=142 (2.8%); Sample not processed, n=102 (2%); Wrong container type, n=94 (1.9%); Sample delivery error (no date of collection), n=105 (2.1%); Sample unlabelled, n=97 (1.9%); Sample volume error, n=2 (0.04%); Laboratory accident, n=1 (0.02%); Other, n=97 (1.9%)

[c] unclear what proportion due to each problem. Not all problems inherent to FIT test, e.g., missing clinical details was important to study, but not to the processing of FIT in clinical care.

4.3.14.2 “Time to” outcomes

Eight studies (nine publications)[47, 49, 52, 63, 64, 69, 72, 78, 84] reported other outcome data listed in the NICE scope. It should be noted that, in accordance with the protocol, data relating to these outcomes were only sought from studies that were included in the diagnostic test accuracy review. The data are summarised in Table 29.

“Time to” outcomes: Six studies[47, 49, 69, 72, 78, 84] reported data on the time to different points within the diagnostic pathway for patients receiving FIT. Amongst four studies[47, 49, 69, 78] relating to single FIT, one[47] reported time to return FIT result (median 7 days (IQR 4–11 days)), another[49] reported time to analysis of FIT (averaged 10.1 days), one[78] reported time to investigation (median 21 (IQR11-43) days) and one[69] reported time to diagnosis (median 59 days, range 8–114 days). One of these also reported that 12 of the 15 patients who had a negative FIT but who had CRC were referred within 2 months, nine of whom were diagnosed within 2 months, and that the median time to diagnosis for the 15 patients was 51 days (IQR 36.5–174.5 days), indicating some patients have a relatively long delay to diagnosis. Another study[72] using single FIT reported a number of outcomes (see Table 29) for patients who tested negative by FIT (in this study the threshold was <20 µg/g), but who were eventually diagnosed with CRC. Three categories were reported, FIT<4 µg/g, FIT 4-9.9 µg/g and FIT 10-19.9 µg/g. Median time to diagnosis was <90 days in all categories, though the IQR was as high as 456.5 in the <4 µg/g subgroup and time to diagnosis was extremely long (>1000 days) for a minority of patients and especially in those with FIT<10 µg/g. This study also reported stage at diagnosis for those with missed diagnoses, which are difficult to interpret without comparative data. This study also reported diagnoses in those who failed to return their FIT, and this was 1%.

Two studies[78, 84] reported time to outcomes for Dual FIT. One[78] reported a small increase in the median number of days to investigation for dual FIT (median 26[a] (IQR 17-45)) vs single FIT (median 21 (IQR11-43) days, P<0.050). The other study[84] reported a median 6 days (IQR 5–8) interval between FIT samples.

Other outcomes: One study[63, 64] reported a number of outcomes after introducing FIT into their diagnostic pathway using a threshold of 10 µg/g (see Table 29). Notably, they report a 9.2% reduction in referrals to colorectal services from 4303 in previous year to 3905 after the introduction of FIT, and similarly a 24.1% reduction in gastroenterology outpatient referrals from 2796 in previous years to 2121 after the introduction of FIT. They also report one emergency presentation out of 5372 who had FIT.

Table 29: Studies reporting other outcomes listed in the NICE scope

Author, year	Analyser	FIT provided in	N with CRC/ N analysed (%)	"Time to" outcomes	Other outcomes
Bailey 2023b *in press*[72]	OC-Sensor iO	Primary care	561/35,289 (1.6%)	**Time to diagnosis for false negative FITs, median (IQR)** • FIT<4 µg/g, with CRC (n=26): 83.5 days (39.5-456.5), max 1023 days • FIT 4-9.9 µg/g, with CRC (n=37): 83.0 days (44.5 - 192.5), n=3 >1000 days • FIT 10-19.9 µg/g (n=25): 41.0 days (26.5-78.0) • FIT<20 µg/g (n=88): 64.0 (34.5 – 212.5), 23/88 >180 days	**Stage at diagnosis:** In the delayed group, 8 (34.8%) patients had Stage I disease at diagnosis, 4 (17.4%) Stage II, 6 (26.1%) Stage III, 4 (17.4%) Stage I and in 1 cancer staging was unavailable. **CRC in patients who did not return FIT:** 38/3631 (1%) **CRC in patients with repeat test:** 62/6640 (0.9%)
Cama 2022[47]	OC-Sensor iO	Primary care	74/5341 (1.39%)	**Time to return FIT result:** median 7 days (IQR 4–11 days) **Diagnostic delay due to negative FIT (n=15):** • <2 month delay to referral: n=12/15 • <2-month delay in diagnosis: n=9/15 • Time from negative FIT to CRC diagnosis (n=15): median 51 days (IQR 36.5–174.5 days).	
D'Souza 2020a[49]	HM-JACKarc	Secondary care	12/298 (4.03%)	**Time to analysis of FIT:** averaged 10.1 days.	No adverse events were reported from patients undergoing FIT or colonoscopy.

Author, year	Analyser	FIT provided in	N with CRC/ N analysed (%)	"Time to" outcomes	Other outcomes
Georgiou Delisle 2022a[52]	OC-Sensor iO	Primary care	61/4187 (1.46%)	NR	Urgent 2WW referrals: 1438/4187 FITs or 2060/5672 patients presenting to primary care
Gerrard 2023[78]	HM-JACKarc	Secondary care		**Time to investigation:** median 21 (IQR11-43) days	NR
Mowat 2021[64] & 2019[63]	HM JACKarc	Primary care	105/5381 (1.95%)	• NR	• FIT<10 µg/g emergency presentations: n=1/5372 who had FIT • Referred to secondary care: n=2848/5372 • Followed up in primary care (no immediate referral): n=2521/5372 • Triaged to colonoscopy: n=1381/5372 • Triaged to gastroenterology: n=672/5372 • Triaged to sigmoidoscopy: n=462/5372 • Triaged to colonoscopy: n=83/5372 • Triaged to other assessment: n=179/5372 • Routine colonoscopy: n=345/1381 colonoscopy • Urgent colonoscopy: n=617/1381 colonoscopy, of which n=419 for suspected cancer - also reports upgrading and downgrading due to FIT result. • Not referred to colonoscopy after review by gastroenterologist: n=71/5660 • Referrals to colorectal services: 9.2% reduction from 4303 in previous year to 3905 • Gastroenterology outpatient referrals: 24.1% reduction from 2796 in previous years to 2121
Tang 2022[69]	HM-JACKarc			**Time to diagnosis:** median 59 days, range 8–114 days	NR
DUAL FIT					

Author, year	Analyser	FIT provided in	N with CRC/ N analysed (%)	"Time to" outcomes	Other outcomes
Gerrard 2023[78]	HM-JACKarc	Secondary care	88/2637 (3.34%)	**Time to investigation:** median 26[a] (IQR 17-45)	NR
Turvill 2018[84]	HM-JACKarc	Secondary care	27/476 (5.67%)	**Time to laboratory (1st sample):** median 7.7hours (IQR 4.9-16.7) **Time to laboratory (2nd sample):** median 6.6 hours (IQR 4.5-14.5) **Time between samples:** median 6 days (IQR 5–8)	NR

CRC, colorectal cancer; IQR, interquartile range; N, number

[a] P<0.050 versus single FIT

4.3.14.3 Patient perspectives

Articles identified by the searches described earlier (in Section 4.1.6 of this report) were sifted for patient reported outcomes of patient acceptability. Patient reported outcomes sought were patient views on the acceptability of FIT, expressions of patient preference for FIT versus colonoscopy, and the experience of, and satisfaction with, FIT in patients with suspected CRC symptoms.

Two studies were identified that investigated patient acceptability: Georgiou Delisle *et al. 2*022 [34]; and MacLean *et al. 2*022[35] (see Table 30). Both of these studies recruited a subset of patients from studies included in this report as diagnostic test accuracy studies (in Section 4.3 of this report). Georgiou Delisle *et al. 2*022 [34] recruited participants from the NICE FIT study. MacLean *et al. 2*022 recruited participants from the POC FIT study. [35] Both of these studies included UK patients referred under the 2WW pathway with suspected CRC symptoms (population type 4).

Both studies designed surveys for their study, rather than using pre-existing surveys. Both studies used a Likert scale 1–5 Georgiou Delisle *et al. 2*022 [34] designed the survey based on a literature review of previous questionnaires, with input from study authors, experts and a patient panel, and MacLean *et al. 2*022 [35] designed the survey with input from study authors and expert academics.

The themes investigated by the two studies were not the same. Georgiou Delisle *et al. 2*022 [34] investigated feasibility of FIT; patient feelings of faecal aversion towards FIT; knowledge in relation to bowel cancer; and future test. Twenty-one statements were included on the questionnaire. The themes investigated by MacLean *et al. 2*022 [35] were expectations, satisfaction that colonoscopy/CTC would rule out CRC C and satisfaction if their FIT results had meant avoiding colonic investigation, and patient experience. There were five questions in this survey.

Georgiou Delisle *et al. 2*022[34] summarised the results of themes by converting into binary, that is, positive (strongly agree, agree) and non-positive (neutral, disagree, strongly disagree). MacLean *et al. 2*022 reported positive responses in a similar manner, and also reported mode and median scores.[35]

Table 30: Study characteristics of patient acceptability studies

Author, date Study design	Study aim	Population	Sample size	Questionnaire used
Georgiou Delisle *et al.* 2022b[34] Cross-sectional survey by postal questionnaire Subset of the NICE FIT study	To investigate attitudes and perception of FIT in symptomatic patients	From the NICE FIT study. UK Patients referred under the 2WW pathway with suspected CRC symptoms. Patients may or may not have completed FIT or had colonic investigation	Questionnaires sent by post 3760. Questionnaires returned and analysed n=1151 (30.6% completion rate)	Developed for the study. Based on literature review of previous questionnaires, with input from study authors, experts and a patient panel. Likert scale 1–5 21 statements 4 themes: feasibility of FIT; patient feelings of faecal aversion towards FIT; knowledge in relation to bowel cancer; and future test intentions
MacLean *et al.* 2022b[35] Cross-sectional survey by telephone questionnaire Subset of the POC FIT study	To investigate patient opinions of FIT	From the POC FIT study. UK Patients referred under the 2WW pathway with suspected CRC symptoms. All had both FIT and colonic investigation.	Contacted by telephone 117. Answered survey n=109 (93% completion rate)	Developed for the study. Design by study authors and expert academics. Likert scale 1–5 5 questions, themes: expectations; satisfaction that colonoscopy/CTC would rule out CRC and satisfaction if their FIT results had meant avoiding colonic investigation patient experience

Study results

Georgiou Delisle *et al. 2*022[34] sent out 3760 questionnaires, and 1151 (30.6%) were returned and analysed, whereas MacLean *et al. 2*022 [35] contacted 117 people of whom 109 (93%) completed the survey. The difference in response rates can be explained as the Georgiou Delisle *et al. 2*022 [34]study posted questionnaires alongside a FIT kit, whereas MacLean *et al. 2*022 telephoned participants who had already engaged with services in both returning a FIT and undergoing colonic investigation, and had fewer questions.

Table 31: Patient characteristics of patient acceptability studies

Author, date	Population	Patient age (years)	Patient sex	Ethnicity
Georgiou Delisle *et al. 2*022b[34] From the NICE FIT study.	n=1151 Patients completing FIT alongside survey 99.2% Patients with prior stool test 71.7% Unclear how many had prior experience of colonic investigation; survey sent with FIT prior to colonic investigation (if needed)	Mean 65 25-39 2.4% 40-64 39.7% 65+ 57.8%	Male 45.4% Female 54.6%	White 88.0% Non-white 12.0%
MacLean *et al. 2*022b [35] From the POC FIT study.	n=109 Patients completing FIT prior to survey 100% Patients with colonic investigation prior to survey 100% (colonoscopy 46.8%; CT colonography 45.9%; flexible sigmoidoscopy 7.3%)	Age 20–39 1.8% 40–59 25.7% 60–79 65.1% 80+ 7.3%	Male 43.1% Female 56.9%	NR

Patient demographics were similar across studies in terms of age, 57.8% aged 65 or over in Georgiou Delisle *et al. 2*022 [34] and 72.4% aged 60 or over in MacLean *et al. 2*022, [35] and sex, 54.5% female in Georgiou Delisle *et al. 2*022 56.9% female in MacLean *et al. 2*022 (see Table 31).

Both studies addressed the usability of FIT (see Table **32** and Table 33). In the Georgiou Delisle *et al. 2*022 study,[34] 95.9% patients gave a positive response (agreed or strongly agreed) that the device was easy to use, and 90.2% that the sample was easy to collect. In the MacLean *et al. 2*022 study,[35] 88% gave a positive response to ease of use of the sampling device. Georgiou Delisle *et al. 2*022[34] also found that 96.3% of patients found the instructions understandable.

Although there were patient feelings of faecal aversion, these were found to be able to be overcome by the patients of the Georgiou Delisle *et al. 2*022 study,[34] with 79.2% disagreeing that it was difficult to overcome embarrassment, 77.0% overcoming disgust, and 76.3% disagreeing that collecting a stool sample for FIT is unhygienic. When asked if they'd prefer FIT to colonoscopy, 78.1% of patients agreed or strongly agreed, in the Georgiou Delisle *et al. 2*022 study.[34] 95.9% agreed/strongly agreed that they would use FIT again in the future.[34]

Table 32: Patient acceptability results, Georgiou Delisle study[34]

Author, date	Theme measured	Results, % giving positive answer
Georgiou Delisle *et al. 2*022b From the NICE FIT study	Feasibility of FIT	Instructions understandable 96.3% (95% CI = 95.1% to 97.3%) Easy to use device 95.9% (95% CI = 94.6% to 96.9%) Would return FIT by post (rather than via GP) 90.5% (95% CI = 88.6% to 92.0%) Straightforward to collect 90.2% (95% CI = 88.3% to 91.8%) Prefer FIT to colonoscopy 78.1% (95% CI = 75.6% to 80.4%)
Georgiou Delisle *et al.* [34] From the NICE FIT study	Patient feelings of faecal aversion towards FIT	Could overcome embarrassment 79.2% (95% CI = 76.7% to 81.4%) Could overcome disgust 77.0% (95% CI = 74.9% to 79.4%) FIT wasn't unhygienic 76.3% (95% CI = 73.7% to 78.6%)
Georgiou Delisle *et al. 2*022b [34] From the NICE FIT study	Knowledge in relation to bowel cancer	Optimistic about cure if detected early 93.0% (95% CI = 91.4% to 94.4%) Worried about getting CRC 78.0% (95% CI = 75.5% to 80.4%) Thought having family history of CRC carried increased risk 75.1% (95% CI = 72.5% to 77.5%)
Georgiou Delisle *et al.*2022b [34] From the NICE FIT study	Future test intentions	Understood purpose of FIT 98.2% (95% CI = 97.3% to 98.9%) FIT's ability to detect cancer important deciding factor 97.3% (95% CI = 96.1% to 98.1%) Would use FIT again 95.9% (95% CI = 94.9% to 96.9%) Felt my future health influences my behaviour today 93.5% (95% CI = 91.9% to 94.8%)

Table 33: Table Patient acceptability results, MacLean study

Author, date	Theme measured	Results, Mode, Median; Positive response
MacLean *et*	Expectations	How much expected to be referred to colonic investigation (1 least expected, 5 most expected)

Author, date	Theme measured	Results, Mode, Median; Positive response
*al.*2022b [35] From the POC FIT study		Mode 5, Median 4; Positive response 60%
MacLean *et al.*2022b [35] From the POC FIT study	Satisfaction	Satisfaction that colonic investigation could rule out CRC (1 completely unsatisfied, 5 completely satisfied) Mode 5, Median 5; Positive response 93% If FIT negative, satisfaction to not undergo colonic investigation (1 completely unsatisfied, 5 completely satisfied) Mode 5, Median 4; Positive response 51%
MacLean *et al.*2022b [35] From the POC FIT study	Patient experience	Ease of use of stool sampling device (1 very difficult, 5 very easy) Mode 5, Median 5; Positive response (easy) 88% Ease of colonic investigation (1 very difficult, 5 very easy) Mode 5, Median 4; Positive response (easy) 78%

MacLean *et al.*2022b asked about satisfaction of clinical outcome to rule out CRC, and found that 51% of patients were satisfied/completely satisfied that if their FIT was negative they need not undergo colonic investigations.[35] 14.6% were neutral on this question, and 32.1% were unsatisfied/completely unsatisfied with not being referred for colonic investigation.[35]

Although the questions were asked in a different way, it appears that a higher proportion of patients in the Georgiou Delisle *et al.* 2022 study[34] have confidence in FIT, with 78.1% preferring it to colonoscopy, whereas only 51% of patients from the MacLean *et al.* 2022b study[35] would be satisfied that negative FIT could rule out the need for colonic investigation. The difference in patients could explain this, as all those in the MacLean *et al.* 2022b study[35] had undergone colonic investigation already, whereas patients in the Georgiou Delisle *et al.* 2022 study[34] had not. Equally, the wording of the question may have elicited different responses.

Georgiou Delisle *et al.*2022 analysed responses in relation to covariates.[34] They found that patients were less likely to prefer to use FIT rather than undergo a colonoscopy if they were aged 40-64 (rather than 65 or older) (OR 0.60; 95% CI = 0.43 to 0.84), or lived in London (rather than outside London) (OR 0.50; 95% CI = 0.36 to 0.71). Patients were more likely to say they'd use FIT in the future if they were

white (OR 3.20; 95% CI = 1.32 to 7.75), or had prior experience of stool tests (OR 2.06; 95% CI = 1.03 to 4.13).[34] Patients were more likely to prefer to use FIT rather than undergo a colonoscopy if they returned a FIT that was successfully analysed to produce an f-Hb result (OR 4.32; 95% CI = 1.49 to 12.52), and more likely to say they'd use FIT in the future if they successfully used a FIT (OR 11.08; 95% CI = 2.74 to 44.75). However, only 15 patients did not complete the test successfully, so the small sample size means results should be interpreted with caution.[34] MacLean *et al.*2022b[35] found that those that went on to CT colonography would have been less satisfied using FIT than those that went on to both colonoscopy (median score 3) and sigmoidoscopy (median score 4). Female patients would have been less satisfied using FIT alone (median score 3) compared with males with (median score 4).[35]

In the Georgiou Delisle *et al.* 2022 study, nine patients returned the questionnaire but not the FIT kit. [34] These patients showed similar results to those returning the FIT kit (88.9% found it easy to collect a sample, 88.9% disagreed FIT was unhygienic) however the small sample size means results should be interpreted with caution.

The authors conclusions were that most patients found FIT acceptable,[34 35] but strategies are needed to engage patients with more negative views of FIT [34], and shared decision making of patient and clinician should be considered for patients dissatisfied with relying on FIT results to decide on need for further investigation.[35]

4.3.14.4 Sociodemographic factors

One conference abstract on FIT return rates across demographic subgroups in patients with suspected CRC symptoms was identified by the searches described earlier (in Section 4.1.6).[90] This study was updated by an in press article, Bailey *et al.* 2023[33] which was submitted by one of the authors who was a stakeholder for this assessment.

The Bailey *et al.* 2023[33] study investigated FIT return in UK adult patients with suspected CRC symptoms, with an aim of identifying whether demographics, ethnicity or social deprivation affect FIT return rates.[33] The study was a retrospective review of records within NHS Nottingham and Nottinghamshire Clinical Commissioning Group (CCG) (see Table 34). Data had been recorded prospectively on all adult patients presenting to primary care with suspected CRC symptoms, excluding those with rectal bleeding or mass, who were sent FIT kits by post.[33] Up to 14 days were allowed between FIT request and being defined as non-return.[33] As further FIT requests could be made to non-returners, only the first FIT request for each patient was included in the return rate analysis.[33] Exclusion criteria for the analysis were: rectal bleeding or mass; duplicate request; request from out of area; sampling error; incomplete request; not indicated under 18 years old; incomplete records.[33] Socioeconomic data were derived from 6-digit postcodes using the Index of Deprivation tool.[33]

Table 34: Study characteristics of equity study[33]

Author, date	Study design	Setting	Population	Sample size	Outcome
Bailey *et al. 2*023[33]	Observational, retrospective review of records	FIT as a triage tool in primary care, NHS Nottingham and Nottinghamshire Clinical Commissioning Group, UK. November 2017 to December 2021	Adult patients with suspected CRC symptoms (excluding rectal bleeding/mass)	First FIT requests for 38,920 individual patients	FIT return in symptomatic patients, by demographics, ethnicity and social deprivation

Study results

The results of the study are summarised in Table 35. For the overall population, FIT return rate was 35,289/38,920 (90.7%). Median age was 66 years, 70.1% were white, and more patients were from the least deprived quintile (28.4%) than any of the other socioeconomic quintiles.

The results of the multivariate analysis of non-returns showed that there were differences in return rate for sex, age, ethnicity and level of socioeconomic deprivation.[33] There was a higher return rate for females (91.0%) than for males (90.2%) (by multivariate analysis, odds ratio of non-return for males with reference female OR 1.11, 95%CI 1.03-1.19).[33] There was a higher return rate for patients 65 years or older (91.9%) than for patients under 65 (89.2%) (OR for non-return age 65+ with reference under 65, OR 0.78 95%CI 0.72-0.83). [33] There was a higher return rate for white patients (91.2%); than for Asian (83.8%) (OR 1.82, 95%CI 1.58-2.10); black (86.6%) (OR 1.21, 95%CI 0.98-1.49); and mixed/other ethnic groups (87.2%) (OR 1.29, 95%CI 1.05-1.59). [33] There was a higher return rate for the least socioeconomically deprived quintile (93.6%), than for more socioeconomically deprived groups, with the most socioeconomically deprived quintile having a return rate of 86.3% (OR 2.20 95%CI 1.99-2.43).[33]

Although not an equity study, Georgiou Delisle *et al. 2*022 reported lower rates of return for both questionnaire and FIT from sites in London, than from sites outside London. [34] Response to the questionnaire was higher in older patients, but there were no significant differences for sex or deprivation, however this was data for the questionnaire only, and demographics were not available by FIT return or non-return.[33]

CRC was diagnosed in 599 patients in the Bailey *et al.* study, of whom 561 returned their first FIT, and 38 were first FIT non-returners.[33]

The authors conclusion was that there is a need to find strategies to mitigate the lower FIT return rates in patients with suspected CRC symptoms who are male; aged under 65; from Asian, black or mixed/other ethnic groups; or socioeconomically deprived.[33] Strategies my involve following up after FIT non-return, information provision in a range of languages, and counselling regarding perceived risk of disease and success of treatment.[33]

4.4 Data selected to enter the cost-effectiveness model

From the analyses, the EAG concluded that the assumption that tests should be considered separately was supported by the comparative diagnostic test accuracy studies, one of which notes that more work is required to understand the clinical impact of the use of different tests. The EAG also concluded that it was not necessary to exclude studies potentially enriched by FIT positives, and that it was not necessary to exclude studies that had <90% receiving colonoscopy as the reference standard since estimates were largely similar. Finally, the EAG concluded that since the estimates by population type were similar and had overlapping credible intervals, the same estimate of sensitivity could be assumed for FIT used in all patients presenting to primary care, compared to FIT used in DG30 low-risk patients (i.e., the current care arm of the model, where FIT is only used in DG30 patients). This impact of this assumption for DG30 low-risk patients was tested in a scenario analysis in the economic model (see Section 5.3.5.1).

Table 35: Results of equity study Bailey *et al.* 2023[33]

Demographic variable	Demographic category	Population Number (% of participants in category)	Returned FIT Number (% of participants returned FIT)	Non-returned FIT Number (% of participants non-return)	OR of non-return (95% CI) Multivariate Logistic Regression Analysis[a]
Sex	Female	21800 (56)	19841 (91.0)	1959 (9.0)	Reference
Sex	Male	17112 (44)	15442 (90.2)	1670 (9.8)	**1.11 (1.03-1.19)**
Sex	Unknown	8 (0.0)	6 (0.0)	2 (0.0)	NA
Age	<65 years	18029 (46.3)	16080 (89.2)	1949 (10.8)	Reference
Age	≥65 years	20891 (53.7)	19209 (91.9)	1682 (8.1)	**0.78 (0.72-0.83)**
Age	Unknown	0 (0.0)	0 (0.0)	0 (0.0)	NA
Ethnicity	White	27277 (70.1)	24864 (91.2)	2413 (8.8)	Reference
Ethnicity	Asian	1584 (4.1)	1328 (83.8)	256 (16.2)	**1.82 (1.58-2.10)**
Ethnicity	Black	801 (2.1)	694 (86.6)	107 (13.4)	1.21 (0.98-1.49)
Ethnicity	Mixed/Other	876 (2.3)	764 (87.2)	112 (12.8)	**1.29 (1.05-1.59)**
Ethnicity	Unknown	8382 (21.5)	7639 (91.1)	743 (8.9)	0.99 (0.90-1.08)
Social Deprivation	5th Quintile (least deprived)	11036 (28.4)	10328 (93.6)	708 (6.4)	Reference

Demographic variable	Demographic category	Population Number (% of participants in category)	Returned FIT Number (% of participants returned FIT)	Non-returned FIT Number (% of participants non-return)	OR of non-return (95% CI) Multivariate Logistic Regression Analysis[a]
Social Deprivation	4th Quintile	6278 (16.1)	5808 (92.5)	470 (7.5)	**1.18 (1.04-1.33)**
Social Deprivation	3rd Quintile	6454 (16.6)	5885 (91.2)	569 (8.8)	**1.39 (1.24-1.56)**
Social Deprivation	2nd Quintile	6177 (15.9)	5521 (89.4)	656 (10.6)	**1.68 (1.50-1.87)**
Social Deprivation	1st Quintile (most deprived)	8927 (22.9)	7703 (86.3)	1224 (13.7)	**2.20 (1.99-2.43)**
Social Deprivation	Unknown	48 (0.1)	44 (91.8)	4 (8.2)	1.28 (0.46-3.57)

NA=not applicable. [a] variables in the multivariate logistic regression analyses were gender, age, ethnicity and socioeconomic deprivation. OR, higher numbers reflect higher *non-return rate* (that is, lower return rate), Numbers in bold indicate confidence interval does not cross 1.

5 COST-EFFECTIVENESS

This chapter presents a systematic review of economic analyses of FIT for symptomatic patients suspected of CRC, and the methods and results of a *de novo* health economic model developed by the EAG comparing the different strategies that include FIT.

5.1. Review of existing health economic analyses published

5.1.1. Cost-effectiveness and HRQoL review - methods

Systematic searches were undertaken to identify existing economic evaluations of the use of FIT in people presenting to primary care with symptoms of CRC. Since a systematic review of the literature on this topic had been performed for the previous appraisal of FIT for patients with suspected CRC (NICE DG30),[11] the EAG's searches included only studies relevant to the decision problem which were published since the previous assessment. The main focus of this review was to explore methodological choices made in previous economic evaluations and their potential relevance to the current decision problem and the model being developed by the EAG, rather than to assess the individual results of published economic evaluations.

Search strategy

A comprehensive search was undertaken to systematically identify economic evaluations of FIT in people with symptoms of CRC. A combined search was also performed using similar search strategies to identify HRQoL studies in the relevant population.

Literature searches were undertaken to identify economic evaluations and studies reporting utility estimates in people with symptoms of CRC were undertaken in February 2023 in the following electronic databases:

- MEDLINE(R) and Epub ahead of print, In-Process Citations and Daily Update (Ovid), 1946 to February 22, 2023
- EMBASE (Ovid), 1974 to 2023 Week 07
- EconLit (Ovid), 1886 to February 09, 2023
- The Cochrane Library (via Wiley), Dec 2022 to Feb 23, 2023
- Tufts' CEA Registry (https://research.tufts-nemc.org/cear4/Home.aspx), from 2016 to February 23, 2023

Searches on the Research Papers in Economics (RePEc) (http://repec.org/) database were not carried out due to time and technical constraints; however, the EAG believes that this is likely to have had only a minor impact on the final results of the review.

The search strategies comprised MeSH and Emtree terms and free-text synonyms for FIT (including terms for each FIT brands), colorectal cancer, and (i) economic or (ii) HRQoL with free-text synonyms for 'EQ-5D'. Searches were translated across databases and were not limited by language. Searches were limited to results since 2016, considering the date of the last systematic review of this topic undertaken to inform NICE DG30.[11] As the Cochrane Library had already been searched in December 2022 for the clinical review, an update search was run to identify any new studies added between December 2022 and February 2023.

Methodological study type search filters to identify economic evaluations were applied in MEDLINE and other databases where appropriate, and were based on the NHS EED filter and economic filter by the McMaster University HEDGES team (https://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx). The search strategies are presented in Appendix 1.

All references obtained were imported into reference management software (EndNote® version 20), with their respective bibliographic data and abstracts, where duplicate references were subsequently excluded.

Inclusion and exclusion criteria

Study selection was carried out in two stages, based on titles and abstracts, and full texts. Studies were required to meet the following criteria in order to be considered relevant for inclusion in the review:

- Full economic evaluations comparing interventions for CRC which included FIT;
- The population of the study should include the relevant population included in the final NICE scope (patients presenting to primary care with symptoms suggestive of CRC);
- Published in English;
- Available in full text format (studies which were available in abstract form only were excluded from the review).

Other types of studies or publications (primary studies, in animal, *in vitro* or genetically based studies, letters to the editor or comments), and duplicated studies on the same model were excluded.

Data extraction and quality assessment

For economic studies, data extraction focused on: (i) the indicated population, main results in terms of costs, consequences and the incremental cost-effectiveness of the alternatives compared, and (ii) the modelling methods used, the sources of input parameters, key modelling assumptions and the robustness of the study results. For HRQoL studies, data extraction focused on: (i) the indicated

population, location of study and (ii) main results in reported EQ-5D valuations in patients with different CRC stages, and potentially for other events related to the diagnosis process.

The methodological quality of the included economic studies was assessed using published checklists for economic evaluations and modelling studies.[91]

5.1.2. Cost-effectiveness and HRQoL review results - summary of studies identified

The results of the searches and selection process for economic evaluations and HRQoL studies are presented as a PRISMA flow chart in Figure ***13***.

For economic evaluations, a total of 820 citations were initially identified after exclusion of duplicates, with 792 being excluded at the title and abstract phase of the selection process. Most of the exclusions were of non-economic evaluation studies or economic evaluations undertaken in populations which differ from that described in the NICE scope (e.g., in asymptomatic patients). Twenty-three studies were reviewed at the full text stage; however, none of these included FIT as part of the diagnostic options for symptomatic patients and were deemed not relevant to the decision problem. A list of excluded studies and comments on each exclusion for both reviews are presented in Appendix 10. These consisted of papers with only the abstract provided, editorial papers or comments, study types other than economic evaluations, studies in a different population than patients with symptoms suggestive of CRC presenting to primary care and studies that did not include FIT. Some studies were excluded for more than one reason (the most outstanding were considered for counting purposes). Two additional studies were retrieved from the HRQoL studies' review, and were included in the final review (Westwood *et al.* 2017 and Medina-Lara *et al.* 2020).[92, 93]

For HRQoL studies, a total of 264 citations were initially identified after exclusion of duplicates. At the title and abstract selection phase, 246 papers were excluded, whilst 18 were reviewed in full text. One study could not be retrieved by the EAG and was therefore excluded. None of the studies met the inclusion criteria (Figure ***13***). The main reasons for exclusion were because studies they did not report EQ-5D estimates, they were reported only as abstracts, or because they reflected in a different population to that listed in the NICE scope. Two studies were economic evaluations which were reviewed and included in the review of economic evaluations.

Table 36 and Table ***37*** summarise the two included economic evaluations. Both studies were model-based cost-utility analyses which report the incremental cost per QALY gained for FIT compared with a variety of comparators as part of the diagnostic pathway for people with symptoms of CRC. Both

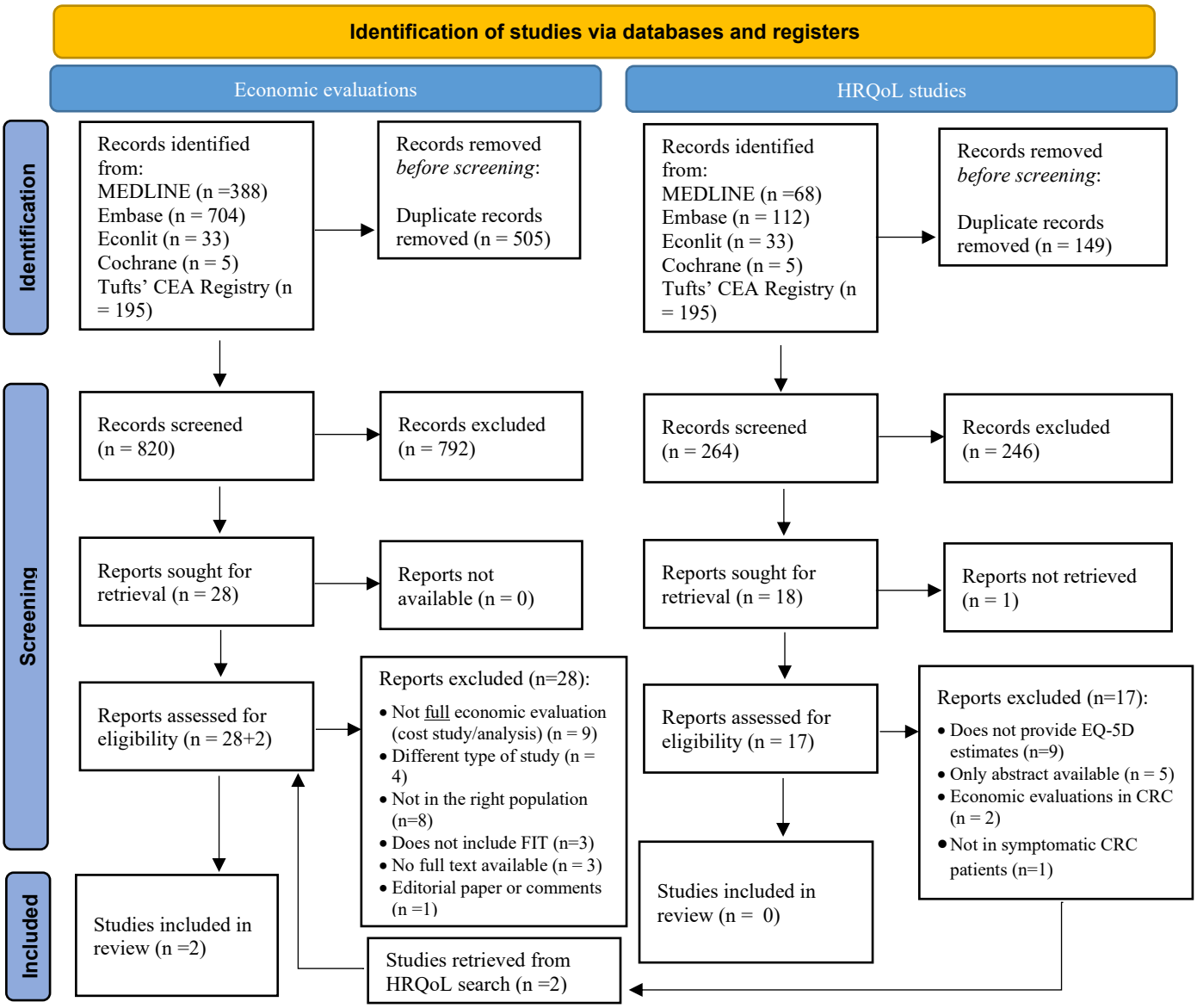
studies were undertaken from the perspective of the National Health Service (NHS) and Personal Social Services (PSS). The models included populations with initial ages ranging from 40 to 70 years.

Both included studies adopted similar general modelling approaches and structures. Both models combined a decision tree containing the diagnostic decision nodes with Markov models which estimate lifetime costs and QALYs for patients with CRC based on states using CRC Duke's staging system, and a two-state model (alive-dead) for patients without underlying CRC. Westwood *et al.* (2017)[92] reported keeping similarities from their structure to previously published model in NICE NG12,[10] whilst Medina-Lara *et al.* provide a comparison of key model characteristics from previous models retrieved from their reviews.[93]

Both models include only CRC patients presenting to primary care who are classified as DG30 low-risk based on NG12/DG30 criteria, and do not include any other underlying lower bowel conditions. The diagnostic component of the model in both studies is based on the prevalence of disease (both assume prevalence of CRC in this population as being 1.5%), and on the accuracy estimates of the tests used for detection of CRC. Both models assume colonoscopy is a perfect diagnostic test, assuming its sensitivity and specificity to be 100%. Both models adopt a lifetime horizon, with cycle lengths ranging from 28 days to one year. A list of assumptions adopted by the models and the sources of their key parameters are presented in Table 36 and Table **37**.

The quality assessment of the included studies is presented in Appendix 10. Considering that the models identified by the review adopted a similar modelling approach, included FIT as the intervention evaluated, and included part of the population considered relevant for this appraisal (they have included only DG30 low-risk symptomatic patients), both existing models informed the development of the EAG's model.

Figure 13: PRISMA flow diagram, review for economic evaluations and HRQoL studies



CRC – colorectal cancer; EQ-5D - EuroQol- 5 Dimension; FIT - faecal immunochemical test; HRQoL – health-related quality of life

Table 36: Existing economic evaluations – analytic scope

Author	Country	Population	Intervention	Comparator	Population's characteristics	Underlying conditions included	Perspective of analysis	Time horizon	Discount rate
Westwood *et al.* (2017)[92]	UK	Symptomatic people who are at low-risk of CRC (as per NG12 definition) presenting to primary care	FIT (10 µg/g threshold chosen based on optimal threshold for each assay method)	• gFOBTs • no triage (all referred to colonoscopy)	Base-case: initial age = 40 years old; proportion of females = 65% CRC prevalence = 1.5%	CRC only	NHS and PSS	"*Lifetime*"	3.5% for QALYs and costs
Medina-Lara *et al.* (2020)[93]	UK	Symptomatic patients with low-risk for CRC (do not fulfil NICE's NG12 2WW referral criteria) but for whom GP has concerns	Use of diagnostic tools (RAT and QCancer) in combination with FIT	• FIT given to all • Send home/wait • Refer all	Initial age = 70 years old; CRC prevalence = 1.5%	CRC only	NHS	Lifetime (30 years)	3.5% for QALYs and costs

CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; NHS - National Health Service; PSS - Personal Social Services; QALY - quality-adjusted life year;

Table 37: Existing economic evaluations - modelling approach, main assumptions, definition of health states and summary of HRQoL included

Author	Model approach	Cycle length	Model type and states included	Sensitivity and specificity for intervention	Sensitivity and specificity for comparator	Sensitivity and specificity for other tests	Model main assumptions	Sources for survival	Costs included	EQ-5D valuation for health states
Westwood *et al.* (2017)[92]	Combined decision tree and Markov STM	Assumed 1 year time frame for diagnostic model; 1-year cycle length for Markov STM	Based on sens and spec of tests (FIT, FOBT and COL/CTC) and symptoms persistance; two Markov STMs to estimate long-term effects, based on CRC Duke's stages (A to D) where patients may stay in current health state, progress to next worst stage or die (from CRC or another cause); and alive and dead states for people without underlying CRC	Base-case scenario (FIT 10 µg/g faeces threshold, single sample): OC Sensor sens 92.1% (95% CI 86.9% - 95.3%); spec 85.8% (95% CI 78.3% - 91.0%) HM-JACKarc sens 100% (95% CI 71.5% - 100%); spec 76.6% (95% CI 72.6% - 80.3%)	FOBT: sens 50% (95% CI 15.0% - 85.0%) spec 88% (85.0% - 89.0%)	COL used as reference standard and assumed sens and spec for detection of CRC to be 100%	Diagnostic model: - patients whose symptoms do not persist assumed not to have CRC; - FN gFOBT or FIT patients whose symptoms persisted were assumed to receive COL and be diagnosed within 1 year and higher probability of progressing to a worse cancer state due to the delay in diagnosis - only those patients with a negative test result who symptoms do not persist do not receive COL/CTC	Patients with CRC: 15-year predicted survival data from NG12; CRC-related mortality assumed constant after year 15 Patients without CRC: UK life tables	- Initial and follow-up investigations - Staging - Lifetime treatment for CRC - Drug costs - Clinical visits and other resources required Costs taken from NG12 CRC treatment lifetime costs from Tappenden *et al.* (2007), inflated to 2015 prices HCHS index	Utilities for CRC stages based on Ness *et al.* 1999 Values used for Dukes' stages: A = 0.74; B = 0.70; C = 0.50; D = 0.25 Population without CRC: sex- and age-related utilities for every cycle from Kind *et al.* (1999)

Author	Model approach	Cycle length	Model type and states included	Sensitivity and specificity for intervention	Sensitivity and specificity for comparator	Sensitivity and specificity for other tests	Model main assumptions	Sources for survival	Costs included	EQ-5D valuation for health states
							Patients with CRC: 1-year cycle length assumed to capture the probability of progression to next worst stage or die for treated and untreated patients Patients without CRC: difference in costs only due to tests and COL/CTC; difference in survival due to COL/CTC			
Medina-Lara *et al.* (2020)[93]	Combined decision tree and Markov STM	28-day cycles	Based on prevalence of CRC, sens and spec of strategies (diagnostic tool, FIT, send home/wait) Markov STM based on	QCancer: (Hippsley-Cox 2012) Sens 0.610; spec 0.910 RAT: Sens 0.69; Spec 0.77 (Hamilton *et al.* 2005)	FIT (threshold of 20 µg/g, Murphy *et al.* 2017): Sens 0.526; Spec for 50–69 years old: 0.988	COL sens and spec =1.0	In the intervention, patients with threshold score above 35 would be directly referred, whilst those with lower values receive FIT	CRC mortality: exponential function fitted to digitised KM curves by stage at diagnosis from NCRAS HR for untreated CRC	FIT, GP visits, COL, COL AEs, heath-stage lifetime treatment costs	Age and sex-matched utilities from Ara and Brazier 2010; CRC Dukes'

Author	Model approach	Cycle length	Model type and states included	Sensitivity and specificity for intervention	Sensitivity and specificity for comparator	Sensitivity and specificity for other tests	Model main assumptions	Sources for survival	Costs included	EQ-5D valuation for health states
			diagnosis status (diagnosed or undiagnosed) and CRC Duke's stages (A to D) where undiagnosed patients may stay in current health state, be diagnosed (via the diagnostic decision model) at their current stage, progress to next worst stage or die (from CRC or another cause); and alive and dead states for people without underlying CRC		Spec for ≥ 70 years old: 0.963		(threshold of 20 µg/g); Sens and spec of tests are assumed to be independent; accuracy of diagnostic tests is assumed independent of disease stage at presentation Model allows for partial adherence to the diagnostic protocol; COL sens and spec assumed to be 1 CRC patients who remain undiagnosed after first presentation will have repeated GP visits until diagnosis or death;	from Liu *et al.* 2014;		stages from Ness *et al.*1999: A = 0.74; B = 0.70; C = 0.50; D = 0.25

Author	Model approach	Cycle length	Model type and states included	Sensitivity and specificity for intervention	Sensitivity and specificity for comparator	Sensitivity and specificity for other tests	Model main assumptions	Sources for survival	Costs included	EQ-5D valuation for health states
							strategies' sens determines number of visits before referral; Impact of delays in referral and diagnosis adapted from Whyte *et al.* using data from Tappenden *et al.*; Disease progression rates from Tappenden *et al.*, estimated for asymptomatic patients is assumed to apply to symptomatic population			

CTC – computed tomography colonography; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; NCRAS - National Cancer Registration and Analysis Service; sens – sensitivity; spec – specificity; STM - state transition model

5.2. Review and critical appraisal of economic analyses provided by test manufacturers

No economic analyses were provided to the EAG by the manufacturers of the FIT tests.

5.3. Independent economic evaluation

5.3.1. Scope of the EAG economic analysis

As part of this assessment, the EAG developed a *de novo* model programmed in Microsoft Excel. ® The model compares different diagnostic strategies that include quantitative FIT in a primary care setting for people with symptoms of CRC. The model assesses the health outcomes and costs associated with each strategy over a lifetime horizon from the perspective of the UK NHS and PSS. The scope of the EAG model is summarised in Table 38.

Table 38: Scope of the EAG economic analysis

Population	Adults presenting to primary care with gastrointestinal symptoms or signs indicating a risk of CRC, excluding people with 'bypass symptoms', such as rectal or anal mass, or anal ulceration
Interventions being compared	Three different sets of interventions that include the use of quantitative FIT in primary care were compared, two of them explored a range of different thresholds. These include: • FIT for all patients using one threshold (t) in µg/g to determine referral decisions • FIT for all patients using two thresholds (t_{high} and t_{low}) in µg/g to determine referral decisions • NICE currently recommended diagnostic pathway, with NG12 high/medium-risk patients being directly referred to the urgent suspected cancer pathway and FIT being offered only to DG30 low-risk patients using a threshold of 10 µg/g to determine referral decisions (as defined by DG30 and NG12)[10, 11]
Primary health economic outcome	• Incremental cost per QALY gained • Net monetary benefit (NMB)
Perspective	NHS and PSS
Time horizon	Lifetime (36 years)
Discount rate	3.5% per annum for health outcomes and costs
Price year (currency)	2021/2022 (£)

CRC – colorectal cancer; EAG – Evidence Assessment Group; FIT - NMB - Net monetary benefit; QALY – quality adjusted life year; NHS – National Health Service; PSS – Personal Social Services; t – threshold; thigh – higher threshold; tlow – lower threshold;

Population

Adults presenting for the first time at a general practitioner (GP) surgery (primary care) with signs and symptoms that might be suggestive of CRC. The population in the model includes patients in both the high/medium-risk and low-risk groups defined in NICE DG30 and NG12.[10, 11] This population excludes patients defined in the NICE scope as having 'bypass symptoms' (very high-risk symptoms: rectal or anal mass, or anal ulceration) who are assumed to be directly referred to secondary care as an urgent suspected cancer referral.

Interventions being compared

The model compares three different sets of interventions that include the use of quantitative FIT in primary care as part of testing strategies to guide diagnostic and clinical management of patients with suspected CRC. These are:

- **Intervention 1**: FIT offered to all patients with a single threshold (t µg/g) used to determine subsequent referral decisions, with the range of FIT thresholds considered being determined by the evidence synthesis (see Section 4.3).
- **Intervention 2:** FIT offered to all patients with pairs of FIT thresholds (t_{high} and t_{low} µg/g) used to determine subsequent referral decisions, with the selection of threshold pairs being determined by the output of the evidence synthesis and clinical opinion from the EAG's advisors.
- **Intervention 3:** Use of NICE current recommendations as defined in DG30 and NG12,[10, 11] with all high/medium-risk patients being directly referred to the urgent suspect cancer referral (hereafter referred to as the 2WW pathway) and DG30 low-risk patients being offered a FIT with subsequent referral decisions for this group being based on a FIT threshold of 10 µg/g.

Perspective, time horizon and discount rate

The economic analysis was undertaken from the perspective of the NHS and PSS considering a lifetime horizon. Unit costs were valued at 2021/22 prices expressed in British pounds sterling (£). Health outcomes and costs were discounted at a rate of 3.5% per annum as recommended by NICE.[94] The model assesses the cost-effectiveness of a range of FIT strategies using incremental cost-effectiveness ratios (ICERs) which are reported in terms of the cost per QALY gained for the intervention strategies versus the strategy that reflects current NICE recommendations. The model also assessed net monetary benefit (NMB) and other outcomes of interest listed in Section 3.3.7.

5.3.2. Conceptualisation of the model

In order to develop a better understanding of the diagnostic pathways of patients with symptoms and signs of suspected CRC, the EAG engaged with multiple clinical advisors before starting the conceptualisation of the model. The EAG's clinical experts included ten health care professionals, which included a mix of academics, GPs, gastrointestinal (GI) consultants and surgeons, registrars and biochemists, all with experience in CRC. A questionnaire was sent to all advisors, with seven replies received. The questions and a summary of the responses from the clinical experts is provided in Appendix 11. Their responses were used to inform some of the parameters of the model where there was insufficient information available from other evidence sources, including the EAG's review of existing economic models and targeted reviews undertaken to populate model parameters.

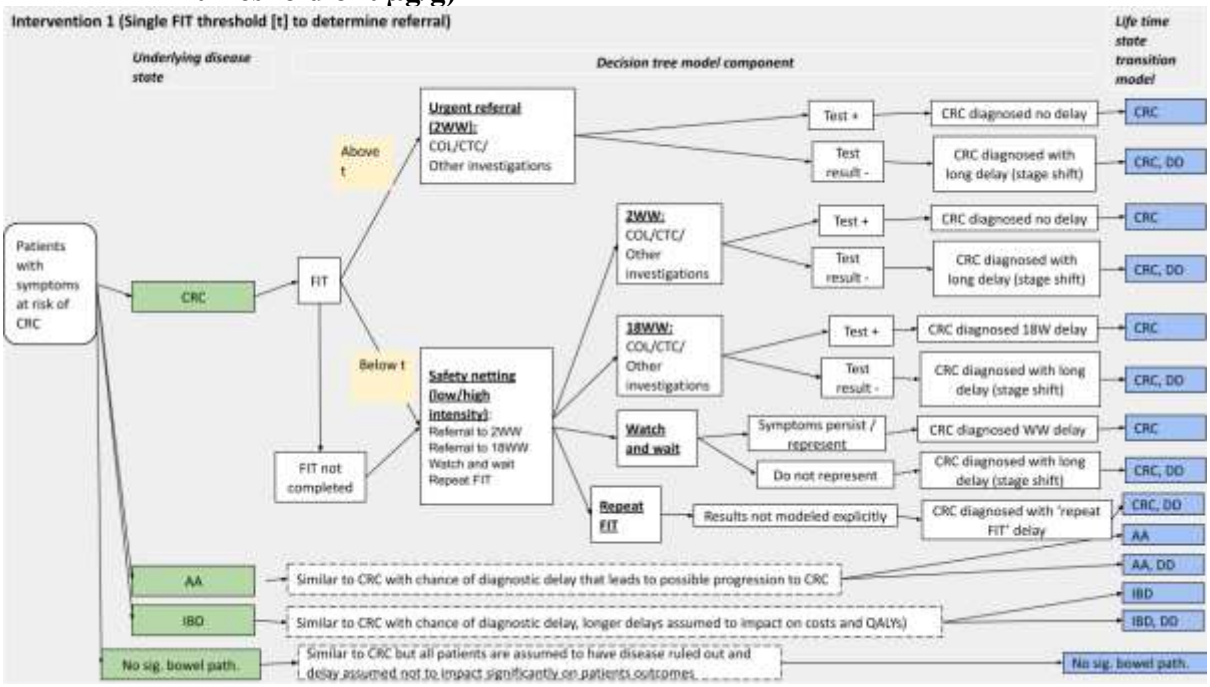
5.3.3. Model structure

The general structure of the EAG's economic model follows a similar approach to that used in the NICE NG12 and DG30 appraisals,[10, 11, 92] and is also broadly consistent with a study identified within the review of published economic evaluations.[93] The model is based on a hybrid structure with a decision tree used to model the diagnostic pathways for a cohort of patients presenting to primary care with symptoms which indicate a risk of CRC. The model is structured to capture the results of investigations reflecting the diagnosis of CRC, but also of AA, and IBD. The decision tree component of the model has a short time horizon which reflects the assumption that the whole diagnostic pathway will cover the period of time between the patient's initial presentation to primary care to confirmation of their diagnosis. Schematic representations of the decision tree part of the model are shown in Figure 14,

Figure 15 and Figure 16 (for Interventions 1, 2 and 3, respectively). The decision tree is followed by state-transition models that estimates lifetime costs, life years and QALYs for people according to their underlying disease state (

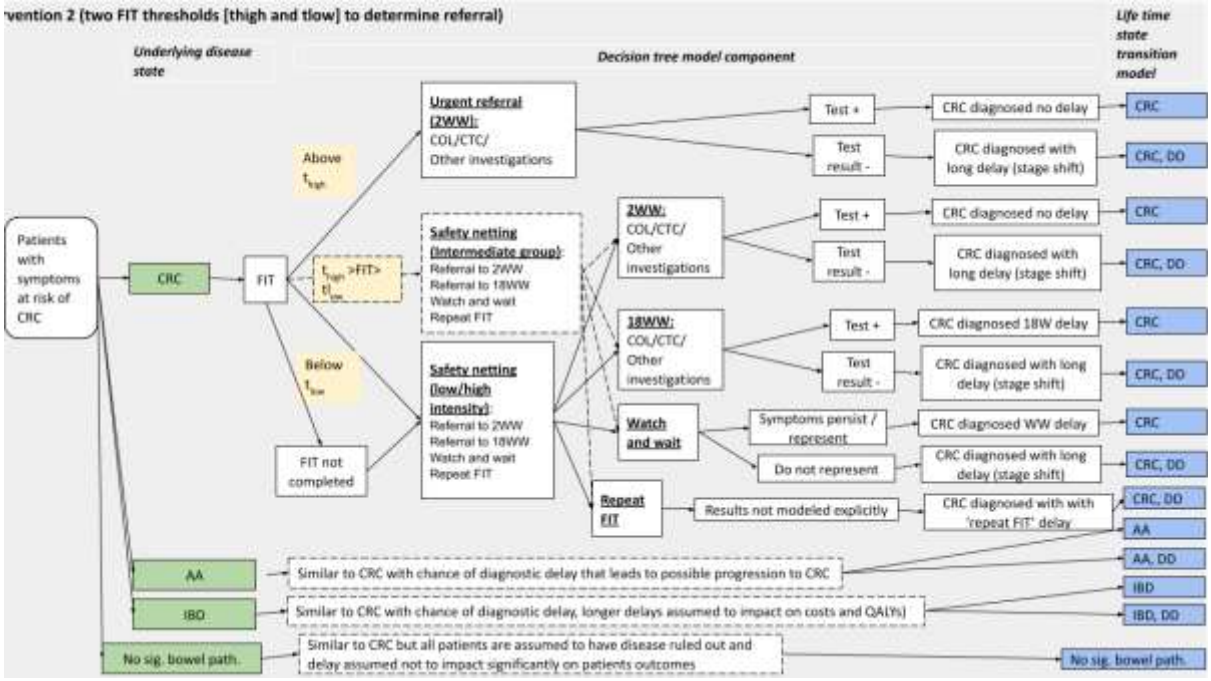
Figure **17**). The model logic is described in Sections 5.3.3.1 and5.3.3.2, whilst the assumptions and sources of parameters used are detailed in Sections 5.3.3.3 and 5.3.4, respectively.

Figure 14: EAG model - decision tree structure, Intervention 1 (FIT with one threshold of t µg/g)



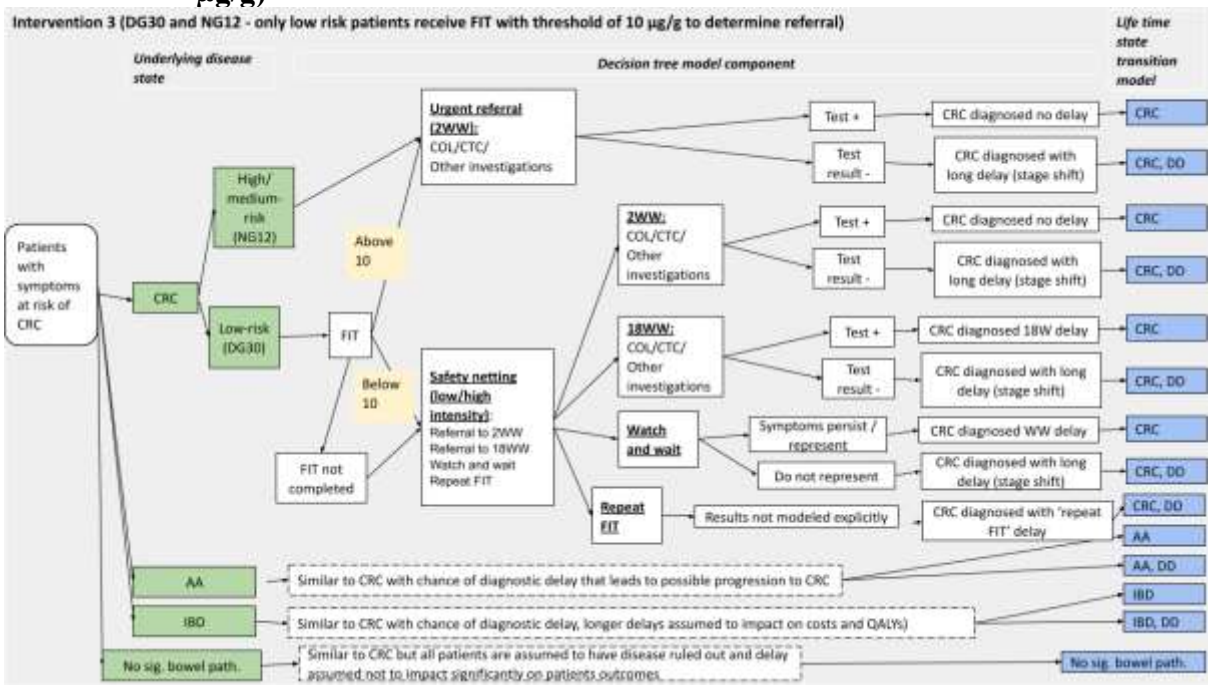
2WW – two week wait; 18W – 18 weeks; 18WW – 18 week wait; AA – advanced adenomas; CTC – computed tomography colonography; COL - colonoscopy; CRC – colorectal cancer; DD – delayed diagnosis (see section 5.3.4.8); FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; t – threshold; WW – watch and wait conduct.

Figure 15: EAG model - decision tree structure, Intervention 2 (FIT with two thresholds of t_{high} and t_{low} µg/g)



2WW – two week wait; 18W – 18 weeks; 18WW – 18 week wait; AA – advanced adenomas; CTC – computed tomography colonography; COL - colonoscopy; CRC – colorectal cancer; DD – delayed diagnosis (see section 5.3.4.8); FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; t_{high} – higher threshold; t_{low} – lower threshold; WW – watch and wait conduct.

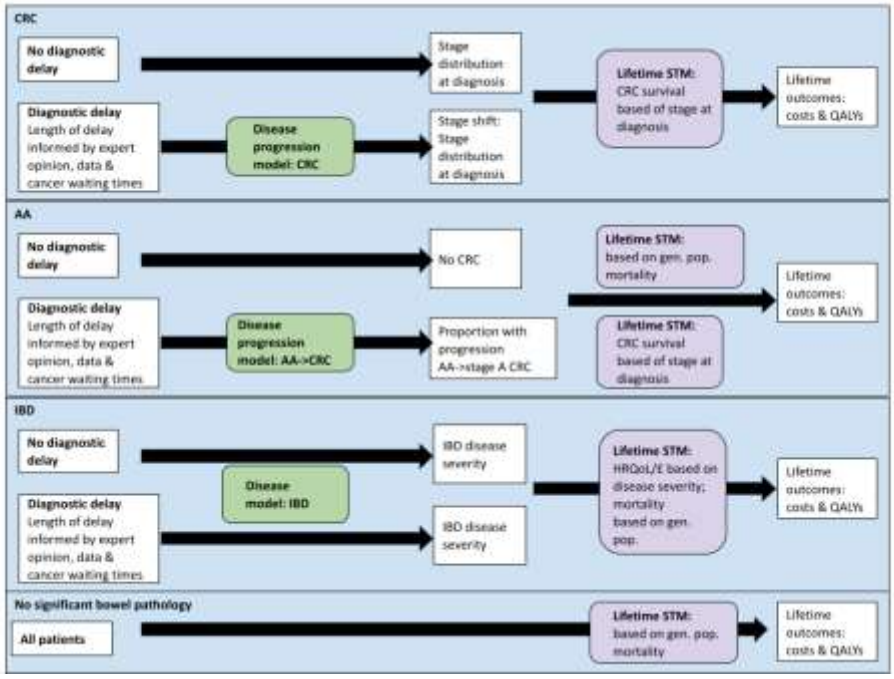
Figure 16: EAG model - decision tree structure, Intervention 3 (DG30 and NG12 recommendations, FIT for DG30 low-risk patients with threshold of 10 µg/g)



2WW – two week wait; 18W – 18 weeks; 18WW – 18 week wait; AA – advanced adenomas; CTC – computed tomography colonography; COL - colonoscopy; CRC – colorectal cancer; DD – delayed diagnosis (see section 5.3.4.8); FIT –

quantitative faecal immunochemical test; High-risk, High/medium-risk; IBD - inflammatory bowel disease; WW – watch and wait conduct.

Figure 17: EAG model - state transition model



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

5.3.3.1. Short term decision-tree component of the model

The EAG model simulates the diagnostic management of patients who present to primary care with symptoms suggestive of CRC, and includes patients who might have underlying CRC, IBD or AAs. IBD and AAs were also included as underlying disease states as they were considered to reflect other significant bowel pathologies which are relevant to the decision problem according to clinical opinion from EAG's advisors. Patients enter the decision-tree component of the model according to their underlying disease, based on the prevalence rates for CRC, AAs and IBD or with no bowel disease and are assumed to enter the model aged 64 years. Under Interventions 1 and 2, all patients are invited to complete a FIT, whilst under Intervention 3, only those patients in the DG30 low-risk group[11] are invited to complete the test, with the NG12 high/medium-risk group patients being directly referred to the 2WW pathway in secondary care.[10]

Patients who complete a FIT and have a result above the threshold (t µg/g in Intervention 1, t_{high} µg/g in Intervention 2 or 10 µg/g in Intervention 3) are directly referred to the 2WW pathway. Patients who do not complete the FIT or whose test result lies below t, t_{low} or 10 µg/g are assumed to receive 'safety netting'. For the purpose of this model, safety netting was defined as being the possible subsequent diagnostic decisions made by the health care professional in primary care following a 'negative'/low FIT result, and includes patients receiving one of the following options: (i) referral to the 2WW pathway

in secondary care; (ii) referral to the non-urgent referral pathway in secondary care (referred hereafter as the 18 week wait pathway, 18WW); (iii) watch and wait (also known as watchful waiting, which consists of patients being monitored in primary care with symptoms reviewed by the GP or patient re-presentation if symptoms persist or worsen); or (iv) invitation to receive a second FIT. For Intervention 2 only, patients who complete the FIT and receive a result that lies between t_{high} and t_{low} are assumed to follow the 'intermediate group' pathways, which the model defines as the same pathway options available in safety netting with the exception of 'watch and wait', and with a higher proportion of patients being directly referred to 2WW (see more details about the pathways and parameters included in safety netting and the 'intermediate group' pathways in Section 5.3.4.4). The inclusion of the direct referrals to secondary care 2WW and 18WW pathways is intended to reflect patients for whom GPs still have important clinical concerns even after the FIT result is returned and referral for further investigations in secondary care is considered necessary.

Patients who are referred to 2WW and 18WW are assumed to receive diagnostic imaging and other tests at secondary care gastroenterology visits. In particular, patients are assumed to receive one of the following imaging investigations: colonoscopy, CTC, or 'other non-invasive investigations'. The EAG opted to model explicitly only colonoscopy and CTC due to these imaging modalities being considered the most common tests used in lower GI referrals by the EAG's clinical advisors (see Appendix 11) and because these had been included in previous economic evaluations.[92, 93] A third option denoted 'other non-invasive investigations' is also modelled to account for the group of patients who would not receive any invasive imaging investigations due to advanced disease stage, frailty, older age or patient/clinician's choice, and was intended to avoid the need to model such options separately to reduce the complexity of the model. The proportions of patients receiving each of the diagnostic imaging test in secondary care are presented in 5.3.4.5. These estimates were informed by clinical opinion from the EAG's advisors, and are conditional on the type of referral received (2WW or 18WW).

The model assumes that all patients with underlying CRC/IBD/AA at either 2WW or 18WW will eventually result in the diagnosis of the underlying disease but with different times to diagnosis, based on the accuracy of the imaging tests received. Patients receiving a colonoscopy as the first imaging test are assumed not to receive any other confirmatory imaging tests, whilst patients receiving CTC with a positive result are assumed to receive a confirmatory colonoscopy. More details on the tests accuracy parameters are presented in Section 5.3.4.6; the EAG model differs from some previous models available for symptomatic patients in the literature in that it does not assume perfect accuracy of colonoscopy and CTC as the first imaging test. The model assumes that colonoscopy and CTC can also detect AAs and IBD, based on test accuracy data from the different literature sources (see Section 5.3.4.6). The model assumes that cases missed by colonoscopy or CTC will be eventually detected by other diagnostic techniques whilst incurring an associated long delay in diagnosis. Patients undergoing

'other non-invasive investigations' are assumed to have their underlying disease detected by 2WW and 18WW referrals (i.e., the model assumes that the combination of other modalities have perfect accuracy, considering the group of patients they are reserved for); this assumption was a simplification to restrict model complexity. Patients with an underlying status of 'no significant bowel pathology' (hereafter termed as NSBP) were assumed to have no disease detected at lower-GI referral; i.e., that the diagnostic test (or sequence of tests) used following referral has perfect specificity.

Patients in the model who follow the watch and wait pathway are assumed to be followed-up in primary care, and are eventually diagnosed with their underlying condition, with an associated specific delay for this pathway. The model assumes that these patients would be diagnosed either by re-presentation to GP due to persistence or worsening of symptoms of symptoms, or following subsequent presentation to an Accident and Emergency (A&E) department. The model also includes patients who would be invited to receive a repeat FIT in primary care. In the absence of robust data identified in the EAG's review on the accuracy of a repeated FIT, the results of the second test are not modelled explicitly. Instead, the model assumes that a proportion of the patients with underlying bowel disease invited for a repeat FIT are detected via referrals and the remaining are detected after watch and wait, with a mean delay to diagnosis estimated for the overall group based on the time to diagnosis for each group. Patients with NSBP receiving watch and wait or repeat FIT were assumed not to re-present with persistent symptoms and/or to receive the confirmation of their underlying pathology.

After patients receive the diagnosis of their underlying disease of CRC/IBD/AA or of NSBP, they are assumed to move to each corresponding lifetime state transition model according to their true underlying pathology, where the lifetime costs, LYGs and QALYs and the impact of delays in the time to diagnosis for each pathology are estimated.

Impact of capacity limitations on waiting times and diagnostic delays

The EAG's model, considering the limited capacity availability for both referrals and colonoscopies in the UK NHS noted in NICE's scope[22], estimates the impact of the use of alternative FIT thresholds on the number of referrals and colonoscopies undertaken. In the base case analysis, capacity used is assumed to have a linear impact on waiting times (and consequently on the time to diagnosis of each pathway), based on the levels of referrals estimated for Intervention 3 (which corresponds in the model to the current NICE recommendations). For example, if the demand for referrals in Intervention 1 at threshold t_x results in a 10% reduction in the total number of referrals (2WW and 18WW), it is assumed that the time to obtain a diagnosis on 2WW and 18WW pathways would be also reduced by the same proportion. Similarly, increases in referrals above levels experienced in Intervention 3 would lead to an increase in waiting times for referrals and thus an increase in the time to diagnosis modelled for Interventions 1 and 2 for these pathways.

5.3.3.2. Lifetime state transition component of the model

This section describes the approach used to quantify survival, QALYs and costs for the long-term component of the model. The pathways followed in the decision tree component of the model are assumed to impact on time to diagnosis, and as a consequence, on survival, QALYs and costs associated with each pathway followed by patients.

Evidence on the association between time to diagnosis and CRC outcomes is heterogeneous. A systematic review explored the association between shorter times to diagnosis and more favourable outcome and found that although many studies reported no associations, more studies reported a positive, rather than a negative, association.[95] For patients with underlying CRC, the EAG's model assumes that longer delays in diagnosis may result in disease progression prior to diagnosis (stage shift) and thus have negative impacts on survival, HRQoL and costs (see Appendix 12).

For patients with underlying CRC and AAs, estimates of lifetime outcomes in terms of life expectancy, health-related quality of life (HRQoL), and health care costs were generated from a separate model ('Additional time to diagnosis impact on outcomes' model by Whyte *et al.* 2023) by 'additional time to diagnosis' and were incorporated to the EAG's model. Full details of this separate model are provided in Appendix 12; a brief summary is provided below.

For CRC, the Whyte *et al.* (2023) model comprises two components: (1) patient outcomes (LYs, QALYs and costs) according to patient's CRC stage at diagnosis and age, and (2) estimates of disease progression during the 'additional time to diagnosis' period. Patient outcomes with and without different delays are compared. CRC disease progression is estimated according to the change in stage distribution as a consequence of the additional time to diagnosis.

For AAs, disease progression is estimated based on the proportion of individuals whose AAs transform to CRC during the delay period, with those individuals who progress receiving lifetime outcomes for patients with CRC (when the delay is less than 1 year, patients who progress are assumed to be diagnosed with CRC Duke's Stage A).

The proportion of individuals who experience a stage shift (a worsening in cancer stage during the delay in diagnosis or the progression from AAs to CRC stage) and the differential outcomes by stage, are combined to generate estimates of expected outcomes by additional time to diagnosis. These estimates were integrated in the EAG model by applying these values as payoffs to each branch of the short-term decision tree, generating estimates for lifetime LYGs, QALYs and costs for each diagnostic strategy.

Patients with IBD and no underlying disease were assumed to enter simple state transition models with two states: alive and dead. During each time interval of one year, patients entering these long-term models can either remain alive or die from any cause. For IBD, the model includes only patients with Crohn's disease (CD) and ulcerative colitis (UC), and patients with IBD are assumed to incur specific costs and utilities for the disease that considers the distribution of patients with each of these conditions, and disease severity (see Section 5.3.45.3.4.2). Diagnosis of IBD through the 2WW and 18WW pathways is assumed to be associated with no significant delay that would impact substantially on outcomes, whilst diagnosis through watch and wait, repeat FIT or long delay (false-negative patients eventually diagnosed with a long delay) are assumed to incur in an increased probability of complications as a consequence of the delayed diagnosis for 2 years after diagnosis, and its associated additional costs and QALY losses due to increase in these complications. The lifetime LYGs, QALYs and costs estimates generated for each of the two groups were then integrated to each branch of the short-term decision tree, thereby generating expected estimates of lifetime LYGs, QALYs and costs for each diagnostic strategy.

For patients with NSBP, any additional time to confirmation of the underlying status is assumed to have no impact on lifetime outcomes, and therefore these patients are assumed not to incur in any additional lifetime costs, but LYGs and QALYs are included, with all-cause mortality risks and health utility assumed to follow sex and age-matched estimates for the general population in England.

5.3.3.3. Key EAG model assumptions

The EAG model makes the following structural assumptions:

- The model was designed to reflect the population of patients presenting to primary care with symptoms or signs indicating a risk of CRC. However, the model is also structured to capture incidental findings of other bowel pathologies: AAs and IBD, which includes UC and CD.
- Patients' underlying disease status (CRC, AAs, IBD and NSBP) are assumed to be mutually exclusive and exhaustive. This simplifying assumption, which does not allow more than one relevant disease to be detected at the same time, was necessary due to the structure of the model developed and was anticipated not to have a significant impact on model results. The EAG notes, however, that delays in the diagnosis of AAs could impact on patient outcomes in the long-term model through the possibility of these progressing to CRC (but these patients are still categorised in the model as AA patients).
- Colonoscopy, CTC and 'other non-invasive' investigations in secondary care are assumed to detect only the underlying condition (e.g., they do not allow for false-positive results for IBD in patients with underlying CRC, and *vice versa*).

- Estimates of accuracy of FIT for CRC, AAs and IBD were informed by the EAG's systematic review (see results in Section 4.3), whilst the sensitivity and specificity of colonoscopy and CTC were obtained from the literature,[96-101] with some necessary simplifications/assumptions made in line with previous published models in this disease area.[92, 93, 96]
- The completion rate for FIT was informed by literature.[44] Patients completing FIT are assumed to receive and return the test to their GP practice in a timely manner so that there are no delays in processing.
- The model assumes that a small proportion of FIT samples need to be repeated due to unsuitable samples, but this was assumed not to affect outcomes or time to diagnosis, and was included in the model only in terms of the cost of completed FIT tests. The model also assumes that the maximum number of FIT tests being received by patients was two (the first invitation and the repeat FIT for a proportion of patients).
- The model assumes that all patients who receive a FIT result above t or t_{high} µg/g (Intervention 1 and 2, respectively) will be referred to 2WW, whilst patients who receive a result below t or t_{low} µg/g or do not complete FIT will follow pathways under 'safety netting'. Patients in Intervention 2 whose FIT result lies between t_{low} and t_{high} are assumed to follow 'intermediate group' pathways. Within these pathways, according to clinical judgement by the GP, patients are assumed to be: referred to 2WW or 18WW; monitored and managed in primary care; or offered a repeat FIT.
- Patients referred to 2WW or 18WW are assumed to receive an initial appointment with a gastroenterology consultant, and are offered colonoscopy or CTC as their main imaging investigation; patients are assumed to undertake the test assigned to them with an assumed uptake rate of 100%.
- Patients who receive 'watch and wait' are assumed to incur in an additional 1.9 appointments with the GP. This is intended to reflect the patient receiving timely review or re-presenting to the GP when their symptoms persist or worsen which would subsequently lead to a diagnosis. The majority of patients with underlying disease are assumed to be eventually diagnosed via a referral, and are assumed to incur the costs of a lower bowel referral. A smaller proportion of patients are assumed to only be diagnosed after presenting at A&E; these patients are assumed incur the costs of presenting at A&E.
- Patients who receive repeat FIT are assumed to incur the cost of the additional FIT and an additional 1.9 GP appointments; this is intended to reflect any additional appointments necessary to discuss results and options for further management. These patients are also assumed to incur the costs of 'watch and wait' or a referral to secondary care, based on accuracy estimates for FIT, which is assumed to estimate a weighted mean cost for patients receiving this pathway.

- The model assumes that patients who receive 'watch and wait' or repeat FIT would eventually be diagnosed with their true underlying condition, but their outcomes (costs, LYGs and QALYs) are impacted by delayed diagnosis.
- Where available, accuracy data for CTC and colonoscopy is informed by the literature.[97-102] The specificity of colonoscopy was assumed to be 100% for all underlying pathologies, due to the nature of the test. This assumption is in line with previous models.
- Patients with a positive result for CTC are assumed to receive a confirmatory colonoscopy. This second test is assumed to have a perfect diagnostic accuracy. The EAG notes that this is a simplification to reduce model complexity.
- Patients with NSBP who are referred to 2WW or 18WW are assumed to incur in costs of COL/CTC/other non-invasive investigations and are eventually ruled out from having any of the lower bowel pathologies being modelled. People with NSBP are therefore assumed not to incur any additional lifetime costs.
- The accuracy of the diagnostic tests received as part of 2WW or 18WW referrals is assumed to be independent of the underlying disease stage/severity at patient initial presentation to primary care.
- All patients who receive colonoscopy are assumed to be at risk of adverse events (AEs) associated with this imaging test (including the risk of deaths after perforation). AEs could result in mortality, QALY loss and costs.
- Patients with NSBP or who are diagnosed with IBD who remain alive following the diagnostic decision tree are assumed to have the same mortality risk as the general population in England of the same age and sex.
- Time to diagnosis is assumed to include the time from initial presentation at primary care to a definitive diagnosis. The model assumes that the greatest impact is derived from the additional time a patient with underlying disease has to spend in primary care in 'watch and wait', 'repeat FIT' pathways or by patients who were missed by imaging tests at secondary care to the point at which a correct diagnosis is obtained. Different lengths of time to diagnosis are explored in scenario analyses (see Section 5.3.5.1).
- The Whyte *et al.* (2023) model assumes that a proportion of patients with underlying CRC will experience disease progression to a worse CRC stage, which depends on the length of additional time to diagnosis, and that patients with AAs may develop CRC during the additional time taken to receive diagnosis. It is assumed that patients can transition to a worse state only once per year (see Appendix 12, which contains a report provided in academic-in-confidence), and that no disease-related deaths are incurred during the delay period.
- This Whyte *et al.* (2023) model also assumes that adenomas are asymptomatic, and therefore the impact on health outcomes and costs due to delay in diagnosis is associated only with those patients who progress to CRC in that delay period. Patients in this model diagnosed without

any delays who have not progressed to CRC are assumed to accrue the same lifetime cost and health outcomes as the general population.

- The lifetime costs for CRC patients are assumed to include costs associated with diagnosis and treatment of CRC, including hospitalisations, medications, and palliative care (Appendix 12).
- Patients with IBD or NSBP are assumed to have the same risk of death as the general population. NSBP patients are assumed not to incur in any additional costs from the point of diagnosis and to have the same HRQoL as the general population, whist IBD patients are assumed to incur specific costs for the treatment of the underlying disease, considering disease severity, costs and HRQoL associated with the condition. A significant delay in the IBD diagnosis is assumed to be associated to an increased probability of having disease complications and incur additional costs and QALY losses, which are assumed to be resolved with treatment after 2 years of diagnosis.

5.3.4.Evidence sources used to inform the model parameters

Table 39 summarises the evidence sources used to inform the parameters of the EAG model. The individual parameter values are discussed in further detail in the subsequent sections. In addition to the review of economic evaluations and HRQoL studies, targeted literature searches were undertaken to identify studies to inform the parameters of the EAG's model, such as patients' initial characteristics, CRC stage distribution, accuracy of the imaging tests (COL/CTC), costs, morbidity including AEs associated with colonoscopy and HRQoL. These searches did not constitute a systematic review but followed the principles of NICE Decision Support Unit (DSU) Technical Support Document 13 (TSD13).[103]

Table 39: Evidence sources used in the model

Parameter group	Parameter group	Source
Patients' initial characteristics	Patient initial age	Based on data from Public Health England on CRC prevalence for 2013[104]
	Probability female	D'Souza *et al.* (2020)[17]
Disease prevalence and severity distribution	Disease prevalence for CRC, AAs and IBD	EAG's clinical review and synthesis (See Appendix 13)
	CRC stage distribution at diagnosis	Staging data in England for 2019,[105] for details see description of Whyte *et al.* (2023) model in Appendix 12
	Proportion of high-risk patients in all with suspected symptoms of CRC	D'Souza *et al.* (2021)[75]
	Distribution of patients with UC and CD and by disease severity in IBD	Pasvol *et al.* 2020,[106] Ghosh *et al.* 2015[107]
Tests' characteristics	FIT accuracy (for CRC, IBD and AAs)	EAG's clinical systematic review and analysis (Section 4.3)

Parameter group	Parameter group	Source
	COL and CTC accuracy (for CRC, AAs and IBD)	Thomas *et al.* (2020);[102] Bressler *et al.* (2007);[97] Atkin *et al.* (2013);[98] Lin *et al.* (2016);[99] Martin-Lopez *et al.* (2014);[100] Horsthuis *et al.* (2008);[101] assumption
	'Other non-invasive investigations' accuracy	Assumption
	FIT return rate	Bailey *et al.* (2021)[44]
	Probabilities of having AEs related to COL	Lin *et al.* (2021);[108] Gatto *et al.* (2003)[109]
Health care current pathways	Proportion of patients receiving each of the pathways following FIT result (safety netting and 'intermediate group' results pathways)	Based on EAG's clinical advisors' responses (Section 5.3.4.4 and Appendix 11)
	Proportion of patients receiving each of the imaging investigations in secondary care (2WW and 18WW)	Based on EAG's clinical advisors' responses (Section 5.3.4.5 and Appendix 11)
	Time to diagnosis for each pathway followed (2WW,18WW, watch and wait, repeat FIT, and patients eventually diagnosed with long delay)	Based on EAG's clinical advisors' responses (Section 5.3.4.8 and Appendix 11)
Mortality	CRC and AA mortality	MiMic Bowel model as reported in Thomas *et al.* (2020)[102] and assumption (see Appendix 12)
	IBD mortality	The risk of death of IBD patients were assumed to be the same as the general population
	Other-cause mortality (general population)	National life tables for England 2018-2020 (ONS)[110]
Long term model probabilities of transitions	Probability of transition between CRC states (progressing)	See description of Whyte *et al.* (2023) model in Appendix 12
	Probability of AA progressing to CRC Duke's Stage A	See description of Whyte *et al.* (2023) model in Appendix 12
HRQoL	CRC	See description of Whyte *et al.* (2023) model in Appendix 12
	AAs	See description of Whyte *et al.* (2023) model in Appendix 12
	IBD	Utilities from NICE TA856[111] and TA342;[112] utility multiplier associated with delayed diagnosis from Stark *et al.* (2009);[113] assumption for duration of impact from delayed diagnosis
	General population	Hernandez *et al.* (2022)[114]
	QALY losses due to colonoscopy AEs	Thomas *et al.* (2020);[102] Dorian *et al.* (2014);[115] Ara & Brazier (2011)[116]
Costs (short term model)	FIT costs (tests)	Unit costs for FIT tests, GP appointment and tests from test manufacturers, PSSRU 2022[117] and NHS Reference costs 2021/22[118]; FIT tests that need resampling from MacDonald *et al.* (2022);[59] FIT return rate from Bailey *et al.* (2021)[44]
	Costs of lower GI referrals	Unit costs for appointments and imaging tests from NHS Reference costs 2021/22[118]; proportion of patients receiving CT in 'other non-invasive interventions estimated from clinical opinion
	Costs of watch and wait	Number of additional GP visits estimated by COLOFIT team based on data from Lyratzopoulos *et al.* (2013);[119] proportion of patients who present to AEs from 'Routes to diagnosis for England 2018'[120], clinical visits and A&E attendance unit costs from PSSRU 2022[121] and NHS Reference costs 2021/22[118]

Parameter group	Parameter group	Source
	Costs of repeat FIT	Test manufacturers; MacDonald *et al.* (2022);[59] Bailey *et al.* (2021);[44] number of additional GP visits estimated by COLOFIT team; PSSRU 2022;[121] routes to diagnosis for England 2018[120]
	Costs of AE related to COL	Unit costs from NHS Reference Costs 2021/22[118]
Costs (long term model)	Lifetime treatment costs for CRC	MiMic Bowel model as reported in Thomas *et al.* (2020)[102] (see Appendix 12)
	Lifetime treatment costs for AAs	
	Lifetime treatment costs for IBD	Annual treatment costs from Ghosh *et al.* (2015)[107] uplifted to 2022 using NHSCII index from PSSRU 2022[117]

BNF – British National Formulary; PSSRU – Personal Social Services Research Unit

5.3.4.1. Patients characteristics

Mean age was assumed to be 64 years of age, based on colorectal cancer prevalence for 2013 from Public Health England (rounded down to an integer value).[104] The cohort of patients was assumed to be 54.9% of females, based on D'Souza *et al.* (2020).[17]

5.3.4.2. Disease prevalence and severity/stage distribution

A summary of the data used in the EAG base case model is provided in Table 40. The probabilities of patients having underlying CRC, IBD and AAs were based on the results of the EAG's statistical analysis undertaken as part of the systematic review (see Appendix 13). The model assumes that in Intervention 3 (based on current NICE recommendations DG30 and NG12), the proportion of patients classified as high-risk was 0.537, based on the proportion of CRC patients classified as NG12 in the D'Souza *et al.* (2020) study.[49] This was considered by the EAG team the more appropriate source for this parameter, and was supported by the EAG's clinical advisors (who considered a proportion close to 0.50 reasonable). Whilst the estimates for the CRC, AAs and IBD prevalence for the overall population and high-risk patients were available from the EAG's evidence synthesis, data for DG30 low-risk patients were only available from D'Souza *et al.* (2020).[49] The model was calibrated to ensure the overall prevalence of each lower bowel pathology was the same in all interventions being evaluated.

The stage distribution of patients at CRC diagnosis by Duke's classification used in the Whyte *et al.*(2023) model was informed by staging data for CRC patients in England in 2019 from the National Cancer Registration and Analysis Service (NCRAS).[105] The distribution of patients having UC and CD was based on data from a detailed UK costing study in IBD (Ghosh *et al.* 2015),[107] whist the proportion of patients having UC on the overall population of patients with IBD in the model was taken from Pasvol *et al.* 2020.[106]

Table 40: Parameters related to disease prevalence and disease stage or severity used in the base-case analysis

Parameter	Group	Mean	Source
CRC prevalence	Whole population	0.028	EAG's clinical systematic review and D'Souza *et al.* (2020)[49]
	High-risk group (NG12, Intervention 3 only)	0.044	
	Low-risk group (DG30, Intervention 3 only)	0.010	
IBD prevalence	Whole population	0.027	
	High-risk group (NG12, Intervention 3 only)	0.032	
	Low-risk group (DG30, Intervention 3 only)	0.022	
AAs prevalence	Whole population	0.023	
	High-risk group (NG12, Intervention 3 only)	0.043	
	Low-risk group (DG30, Intervention 3 only)	0.000	
Proportion of patients classified as NG12 high-risk or DG30 low-risk for CRC	High-risk group (NG12, Intervention 3 only)	0.537	D'Souza *et al.* (2020)[49]
	Low-risk group (DG30, Intervention 3 only)	0.463	
CRC stage distribution at disease diagnosis	CRC Stage A	0.196	Staging data in England for 2019 (NCRAS 2021)[105]
	CRC Stage B	0.254	
	CRC Stage C	0.312	
	CRC Stage D	0.238	
IBD disease and disease severity distribution at disease diagnosis	Relative incidence of UC vs CD	0.60	Pasvol *et al.* (2020)[106]
	Proportion of patients in remission with UC	0.50	Ghosh *et al.* (2015)[107]
	Proportion of patients in relapse with UC	0.50	
	Proportion UC relapse mild-moderate	0.80	
	Proportion UC relapse severe	0.20	
	Proportion of patients in remission with CD	0.50	
	Proportion of patients in relapse with CD	0.50	

AA – advanced adenomas; CD – Crohn's disease; CRC – colorectal cancer; IBD – irritable bowel disease; UC – ulcerative colitis.

5.3.4.3. FIT accuracy

Data relating to the accuracy of each of the FIT brands was informed by the EAG's evidence synthesis; a summary of the results of these analyses is presented in Section 4.3.10 of this report. The model used the estimates of sensitivity and specificity for CRC for selected thresholds available (between 2 µg/g and 400 µg/g, selection based on availability of results for each FIT test). For Intervention 1, all of the thresholds were tested individually for those brands with data available. For Intervention 2, due to the excessive number of possible combinations, the values for the thresholds pairs were selected based on the model results for individual thresholds and clinical interest.

For Intervention 2, where the pathways followed by patients were determined by three different groups based on the results of FIT, the results of FIT were calculated as follows:

FIT result $\geq t_{high}$ = sensitivity of FIT for t_{high} (i.e., if t_{high}=100, sensitivity for t_{100})

FIT result $\leq t_{low}$ = 1- sensitivity of FIT for t_{low} (i.e., if t_{low}=10, 1-sensitivity for t_{10})

$t_{low} \leq$ FIT result $\leq t_{high}$ (intermediate group) = 1-(sensitivity t_{high} + sensitivity t_{low})

The value for the FIT return rate of 0.91 was taken from Bailey *et al.* (2021[44]), which was assumed to be the same for the first and second FIT received.

5.3.4.4. Probability of following each of the pathways following FIT result

Patients who receive a FIT result above t, t_{high} or 10 µg/g (in Interventions 1, 2 and 3, respectively), or are classified as being high-risk by NG12 criteria in Intervention 3, are directly referred to the suspected cancer urgent pathway (2WW). However, patients who obtain FIT results below these thresholds or do not complete the test are assumed to follow one of the pathways with two possible groups: safety netting or 'intermediate group' pathways.

Safety netting is defined in this model as the follow up pathways for patients with FIT results below t, t_{low} or 10 µg/g (in Interventions 1, 2 or 3, respectively) or incomplete test, whilst 'intermediate group pathways' is reserved for patients in Intervention 2 for whom the FIT result lies between t_{low} and t_{high}. The model represents safety netting with a proportion of patients following each of four pathways: 2WW; 18WW; watch and wait; or repeat FIT, based on estimates derived from clinical opinion from the EAG's advisors (see Table ***41***). The model explores two different intensities of safety netting pathways (low or high), based on the view that clinical practice is heterogeneous across the country and has been changing in recent years, with the introduction of FIT as part of screening programmes and for triage in symptomatic DG30 low-risk patients.[11] The EAG's model generates results for both options of safety netting.

The pathways for the 'intermediate group' include the same pathway options as for safety netting with the exception of watch and wait and with a higher proportion of patients assumed to be referred to 2WW. This approach is based on the assumption that these patients would be considered having a higher risk of CRC and therefore being referred to secondary care.

Table 41: Proportion of patients receiving each of the management pathways following FIT results*, based on clinical advice provided to the EAG

	Results of FIT			
	Positive FIT (FIT> t, t_{high} or 10 µg/g)	**'Negative' FIT (FIT< t, t_{low} or 10 µg/g) or FIT incomplete**		**'Intermediate' FIT (t_{low} < FIT <t_{high}) - Intervention 2 only**
Proportion following each pathway	Referral to 2WW directly	Safety netting pathways - model base-case (high intensity)	Safety netting pathways - model scenario analysis (low intensity)	'Intermediate group' follow-up pathways
Referral to 2WW	100%	15%	5%	85%
Referral to 18WW	-	25%	10%	10%
Watch and wait	-	40%	75%	0%
Repeat FIT	-	20%	10%	5%

FIT – quantitative faecal immunochemical test; 2WW – two week wait; 18W – 18 weeks; 18WW – 18 week wait; T – single FIT threshold in Intervention 1; t1 – higher FIT threshold in Intervention 2; t2 – lower FIT threshold in Intervention 2.
**In Intervention 3, patients classified as high-risk are referred to 2WW without receiving a FIT.*

In the model, the watch and wait pathway consists of patients' symptoms being reviewed by the GP or patient re-presentation to the GP if symptoms persist or worsen; these strategies are not explicitly modelled separately. The base case model assumes that patients followed-up in this pathway incur an additional 1.9 GP appointments (estimated by a member of the modelling team in the COLOFIT project based on data from Lyratzopoulos *et al.* (2013),[119] obtained via personal communication) and that all patients are eventually diagnosed with their true underlying condition, but a proportion of patients (22.15%) would be only detected at presentation at A&E, based on data from routes to cancer diagnosis in England in 2018 by NHS Digital.[120] Patients following this pathway are assumed to be diagnosed with their underlying condition with a significant delay (see Section 5.3.4.8), which is associated with an increased probability of CRC stage progression or a risk of AAs transforming to CRC during the delay period (see Appendix 12).

Patients invited to receive a second FIT in primary care (repeat FIT) do not have the results of the second test modelled explicitly, due to limited data identified by the EAG's on the accuracy of repeated tests (see Section 4.3.14) and on how the data on Dual FIT is applicable to the population receiving the test in the primary care context. Instead, the model assumes that patients with underlying bowel disease who are invited for repeat FIT would receive the results of the second FIT and would either be diagnosed via a pathway in secondary care or eventually diagnosed by re-presenting to GP with persistent/worsening symptoms which would be associated with a long delay in diagnosis. Therefore, patients in this pathway are assumed to obtain a diagnosis of their underlying condition with a specific delay for this pathway (see Section 5.3.4.8) and to incur additional diagnostic costs (see Section 0)

5.3.4.5. Probability of receiving each imaging test as part of lower-GI referral (2WW and 18WW)

Patients referred to secondary care under 2WW or 18WW are assumed to receive one of the following imaging investigations: colonoscopy, CTC or 'other non-invasive investigations'. The proportions of patients receiving each imaging test in the EAG's base case analysis are presented in Table 42, and were based on opinion from EAG's clinical advisors (see Appendix 11) on the use of imaging tests in referrals in this population (answers for the overall population).

Table 42: Proportion of patients receiving each pathway at their lower GI referral (2WW/18WW)

Investigation received	**2WW**	**18WW**
COL	90%	90%
CTC	7.5%	7.5%
Other/no investigations (e.g., CT or appointment)	2.5%	2.5%

COL – colonoscopy; CTC – computed tomography colonography; 2WW – two week wait; 18WW – 18 week wait.

The EAG notes that responses from clinical advisors and published literature suggest that CTC capacity is currently very restricted in England,[122] and might be one of the reasons why CTC is less frequently used in clinical practice when compared to colonoscopy. The EAG also notes that these proportions vary by age group; however, the estimates used are intended to reflect the usage by the overall population being referred to secondary care with symptoms suggestive of CRC.

5.3.4.6. Accuracy of the imaging tests used in 2WW and 18WW

Data on the accuracy of colonoscopy and CTC received by patients referred to 2WW and 18WW are presented in Table ***43***. The EAG model adopts a similar approach to Thomas *et al.* in the MiMic bowel cancer screening model.[102] Sensitivity estimates for CRC detection by colonoscopy and CTC were based on the studies from Bressler *et al.* (2007)[97] and estimates of the relative risk (RR) for the detection of CRC using CTC rather than colonoscopy from Atkin *et al.* (2013).[98] The EAG notes that the use of these estimates includes the assumption that these imaging tests would have similar performance in symptomatic and asymptomatic patients. Sensitivity for detection of AAs by colonoscopy was based on Martin-Lopez *et al,* (2014),[100] whilst the same approach based on the RRs from Atkin *et al.* (2013)[98] was applied to estimate the sensitivity for AAs by CTC. The specificity of colonoscopy was assumed to be 1.00 for all conditions, given the nature of the test. Specificity estimates for CRC and AAs detection by CTC were taken from Lin *et al.* (2016),[99] and was assumed to be the same for the two conditions. Sensitivity and specificity for detection of IBD by CRC were obtained from Horsthuis *et al.* (2008).[101]

Table 43: Estimates of accuracy for imaging tests used in patients in 2WW and 18WW

Condition	Technology	Parameter	Point estimate	95% CI	Source
CRC	COL	Sensitivity	0.966	0.962 – 0.969	Thomas *et al.* (2020);[102] Bressler *et al.* (2007)[97]
		Specificity	1.000	–	Assumption due to nature of test
	CTC	Sensitivity	0.946	0.606 – 1.473	Thomas *et al.* (2020);[102] Atkin *et al.* (2013)[98]
		Specificity	0.881	0.873 – 0.889	Lin *et al.* (2016)[99]
AAs	COL	Sensitivity	0.925	0.894 – 0.952	Thomas *et al.* (2020);[102] Martin-Lopez *et al.* (2014)[100]
		Specificity	1.000	–	Assumption due to nature of test
	CTC	Sensitivity	0.759	0.465 – 1.218	Thomas *et al.* (2020);[102] Atkin *et al.* (2013)[98]
		Specificity	0.881	0.873 – 0.889	Lin *et al.* (2016),[99] assumption
IBD	COL	Sensitivity	1.000	–	Assumption in line with previous models for CRC symptomatic and asymptomatic patients
		Specificity	1.000	–	
	CTC	Sensitivity	0.843	0.750 – 0.918	Horsthuis *et al.* (2008)[101]
		Specificity	0.951	0.868 – 0.994	

AAs – advanced adenomas; COL – colonoscopy; CRC – colorectal cancer; CTC – CT colonography; IBD - irritable bowel disease.

Patients receiving 'other non-invasive investigations' as part of the diagnostic pathway in 2WW and 18WW are assumed to be diagnosed and no cancer cases are missed (sensitivity and specificity of 1.0 for all conditions), based on the assumption that for this small group of patients with greater frailty and possibly disease severity, a different number of non-invasive diagnostic techniques would be able to detect the patient's underlying condition.

5.3.4.7. Complications associated with colonoscopy

Complications associated with colonoscopy were included in the model for a proportion of patients receiving this imaging test. Patients receiving colonoscopy have a small probability of developing bleeding or perforation of the intestine as a consequence of the procedure (Table 44; these probabilities were based on Lin *et al*. (2021)).[108] Those with perforations can also die as a consequence of the complication; this probability was informed by Gatto *et al.* (2003).[109] Patients having complications from colonoscopy are assumed to incur additional costs and HRQoL losses, which (with exception of death) were assumed to be temporary and resolved without further long-term impacts on their health outcomes after disease diagnosis. Similar to the approach used by Thomas *et al.* (2020),[102] the EAG model includes QALY losses associated with bleeding and non-fatal perforation due to colonoscopy. The utility value for serious bleeding events was taken from Dorian *et al.* (2014)[115] and was assumed to last for 2 weeks, whilst QALY losses due to non-fatal perforation were based on Ara & Brazier

(2011),[116] with the utility value based on the absolute difference in mean EQ-5D score between patients with 'stomach ulcer/abdominal hernia/rupture' who were not affected by the condition and those who were affected by it; this event was assumed to impact on HRQoL for 1 month.

The sources for costs associated with colonoscopies are presented in Section 0. Patients who receive colonoscopy after receiving CTC are assumed to be susceptible to the same AEs and corresponding probabilities as patients receiving colonoscopy as their main imaging investigation.

Table 44: Complications and QALY losses associated with colonoscopy included in the short-term model

Complication	Probability of having an AE	Source	QALY loss	Source
Serious bleeding	0.00175	Lin et *al.* (2021) [108]	0.00579	Thomas *et al.* (2020),[102] based on data from Dorian *et al.* (2014)[115]
Perforation	0.00054	Lin *et al.* (2021) [108]	0.00983	Thomas *et al.* (2020),[102] based on data from Ara & Brazier (2011)[116]
Death by perforation	0.05195	Gatto *et al.* (2003)[109]	*	

COL – colonoscopy
**Patients who die from perforation following a colonoscopy do not incur any QALY losses but are assumed not to receive any QALYs from the point of death.*

5.3.4.8. Time to diagnosis and diagnostic delays

In the model, for patients with an underlying lower bowel pathology (CRC, IBD and AA), the time to diagnosis was assumed to depend on the pathway followed. For example, in the EAG's base-case, patients receiving 2WW were assumed to receive their diagnosis within the period informed by the clinical advisors (see Appendix 11) and not to incur in any delays in their diagnosis. The time to diagnosis necessary for each pathway is presented in Table 45 and was based on clinical input from the EAG's advisors.

The EAG notes that time to diagnosis could be defined in several ways, such as time from symptom onset, time of presentation in primary care or time from referral. In the EAG model, a diagnosis delay is assumed to comprise the additional time to diagnosis compared to average time to diagnosis with a 2WW referral. The model assumes that the estimates of time to diagnosis do not account for small differences in diagnostic interval due to the time for taking the FIT test and receiving its results. For example, patients in Intervention 3 who are referred straight to 2WW would strictly speaking being able to receive a diagnosis faster than patients completing a FIT and in the sequence being referred to 2WW with a positive result, but the model assumes that these small differences would not impact on disease progression and were therefore not considered. Differences in times to diagnosis between FIT and repeat FIT included in the model as shown in Table 45.

Table 45: Estimated diagnostic delays by each type of pathway and diagnostic result†

Pathway followed	Estimated average time to diagnosis for patients with underlying CRC/IBD/AA (weeks)*		
	Model base-case	Scenario analysis 1	Scenario analysis 2
Lower GI referral (2WW) disease diagnosed at referral	2	2	3
Lower GI referral (18WW) disease diagnosed at referral	27 (6 months)	18 (4 months)	54 (1 year)
Lower GI referral (2WW/18WW), disease missed by COL/CTC, patient re-presents with persistent symptoms	78 (1.5 years)	52 (1 year)	157 (3 years)
Watch and wait, patient re-presents with persistent symptoms	59 (1.13 years)	35 (8 months)	104 (2 years)
Repeat FIT (weighted average of subsequent pathways)	38 (8.7 months years)	23 (5.3 months)	69 (1.3 years)

**For patients without underlying disease, the model includes the costs of additional investigations needed for those who have an initial FIT positive test and are referred to 2WW and 18WW and the impact on HRQoL from AEs associated with colonoscopy.*
†The time to diagnosis does not include the time of initial investigations by the GP (initial appointment and FIT)

The EAG also notes that the time to diagnosis for patients receiving a repeat FIT was estimated based on a weighted mean time for the other pathway times and the proportions receiving each of these pathways (high intensity safety netting, see Table ***41***). The model also assumes that all time lengths include the turnaround time required to receive the results of a FIT test and all delays would be in relation to the 2WW pathway.

The estimates presented in Table 45 correspond to a reference point, used to estimate the time to diagnosis for patients in the current scenario in England, which is assumed to correspond to current NICE recommendations. In order to estimate the impact of the introduction of FIT to all symptomatic patients in primary care and the resulting expected impact on waiting times, the EAG included in the model structure the assumption that reductions in the number of referrals to secondary care (2WW and 18WW) would vary by threshold applied and would have a linear impact on the waiting times for these two pathways, e.g. a reduction of 10% in the total referrals as a consequence of a specific threshold would reduce the time to diagnosis for patients receiving 2WW and 18WW in this strategy by the same proportion.

Three different scenarios were explored in the model:

- Base-case scenario. This scenario is intended to reflect the current situation in England
- Scenario 1. This explores a best-case scenario with lower times to diagnosis for all pathways
- Scenario 2. This explores a worst-case scenario where times to diagnosis are increased (see Section 5.3.5.1).

5.3.4.9. Long-term state transition model outcomes

Appendix 12 presents the details on the estimates and source of parameters used in the Whyte *et al.* (2023) model to generate the estimates of lifetime outcomes for CRC and AA patients by 'additional time to diagnosis' in the EAG model.

The risk of death for IBD and NSBP patients were informed by the sex and age-matched mortality estimates for the general population in the England.[110] Utilities for NSBP patients were also assumed to follow the age- and sex-match estimates for the UK population from Hernandez Alava *et al.* (2022).[114] The health utility values used in the model for patients with IBD are summarised in Table 46[113] which is applied to the increase in the proportion of patients having disease complications for two years at the point of diagnosis.

Table *46*; these were estimated based on utility values reported by NICE TA856[111] for UC and TA342[112] for CD (which were based on Woehl *et al.*[123] and the GEMINI II/III studies),[124] and the relative incidence between UC and CD and distribution of patients by disease severity in these conditions as reported in Pasvol *et al.* (2020)[106] and Ghosh *et al.* (2015).[107] The proportions of patients by disease type and severity are summarised in

Table ***47***.

Patients diagnosed with IBD were assumed to incur a QALY loss that was estimated to correspond to the impact of the increased risk of having complications, based on data from Stark *et al.* (2009),[113] which is applied to the increase in the proportion of patients having disease complications for two years at the point of diagnosis.

Table 46: Health utilities applied in the EAG model for IBD patients

Underlying condition	Health state	Mean utility	Source
UC	Active UC	0.41	NICE TA856[111] based on Woehl *et al.* (2008)[123]
	Remission	0.87	
	Response	0.76	
CD	Remission	0.82	NICE TA342[112] based on GEMINI II/III studies[124]
	Moderate-severe	0.57	
Estimate for IBD	All (assumption)	0.75	Estimated based on values for each condition and severity
Utility multiplier for patients having increased		0.73	Stark *et al.* 2009[113]
Increase in IBD complications		0.04	Whyte *et al.* Personal communication

CD – Crohn's disease; IBD – inflammatory bowel disease; UC – ulcerative colitis

Table 47: Proportion of patients by disease severity and type applied in the EAG model for IBD patients

Underlying condition	Proportion of patients	Source
Relative incidence of UC vs CD	0.60	Pasvol *et al.* (2020)[106]
Proportion of patients in remission with UC	0.50	Ghosh *et al.* (2015)[107]
Proportion of patients in relapse with UC	0.50	
Proportion UC relapse mild-moderate	0.80	
Proportion UC relapse severe	0.20	
Proportion of patients in remission with CD	0.50	
Proportion of patients in relapse with CD	0.50	

CD – Crohn's disease; UC – ulcerative colitis

5.3.4.10. Resource use and costs

The model includes the following cost components:

(i) Costs associated with the FIT test (first test)

(ii) Cost of lower GI referrals

(iii) Costs of Watch and Wait

(iv) Costs associated with 'Repeat FIT'

(v) Costs of treating adverse events related to colonoscopy

(vi) Costs associated with treating the underlying conditions (lifetime costs for CRC, IBD and AAs).

FIT costs

The costs for each brand of FIT tests were sourced from information provided to NICE by the manufacturers as part of the appraisal process. The price of each test is provided in Table ***48***. The EAG notes that some of these costs are indicative and may vary depending on some factors, such as type of analyser and methodology employed by laboratories, testing volumes and capacity. The impact of these factors in the costs which will be used in the NHS is unclear.

Table 48: Test costs assumed in EAG analysis

Test[a]	Total cost per test[b]	Comments
NS-Prime	£6.00	Cost per test provided by the company includes the cost of analyser and all consumables.
QuikRead go	£4.40	The manufacturer provided the cost for 50 tests, and the costs of "sampling test", "control quantitative" and "instrument" separately. The total cost per test was estimated based on 50 tests.
HM-JACKarc	£2.31	The cost per test includes rental of the analyser, reagents, consumables, training and servicing, and patient packs.
IDK® Hemoglobin	£6.46	The manufacturer provided the cost for different quantities of tests, and the costs of sampling test and extraction tubes separately for 100 tests. The total

		cost per test was estimated based on the lowest cost per test (96 tests).
IDK® Hemoglobin/Haptoglobin	£6.46	The manufacturer provided the cost for different quantities of tests, and the costs of sampling test and extraction tubes separately for 100 tests. The total cost per test was estimated based on the lowest cost per test (96 tests).
OC Sensor	£4.53	Total cost per test was based on the 'total costs including materials' from DG30.[11] The manufacturer also clarified that this cost included reagent rental of the analyser and that the cost per test is indicative, as it varies by testing volume and methodology employed by the testing laboratory .
FOB Gold	£3.70	Total cost per test based on the midpoint of the range of costs provided by the manufacturer. It is unclear if it includes the costs of other required consumables, or the analyser.

[a] *only tests for which there was diagnostic test accuracy data have been included*
[b] *The EAG notes that it is unclear if the prices provided for NS-Prime, HM-JACKarc and OC Sensor include VAT. For the other brands, the EAG was informed by the manufacturers that they do not include VAT.*

The cost of FIT for patients who complete the test also includes the costs of samples that need to be retaken (proportion of tests that need resampling for technical reasons of 2.1%), based on MacDonald *et al.* (2022).[59] The model assumes that 91% of patients complete the test, based on Bailey *et al.* (2021).[44]

The model also assumes that patients would receive an appointment with the GP to discuss the FIT results (with unit costs for a surgery consultation lasting 9.22 minutes obtained from the PSSRU[121]) and receive additional blood tests, which was assumed to include the costs of one phlebotomy service and one blood count (DAPSS08 and DAPS05 codes from NHS Reference costs 2021/22[118]). Under Intervention 3, the lowest price of all FIT tests is used (price = £2.31). The total cost per patient of the first FIT completed is estimated to be between £46.02 and £50.26, depending on the FIT brand received. Patients who do not return the FIT are assumed to incur only the cost of the test (£2.31 for Intervention 3 and between £2.31 and £6.46 for Interventions 1 and 2).

Costs of lower GI referrals

In the model, patients referred to 2WW or 18WW are assumed to receive, regardless of the type of referral received, an initial appointment with gastroenterologist consultant (£186.48, based on NHS Reference Costs - weighted average of first attendances face-to-face and non-face-to-face with gastroenterologist - Codes 301, WF01B and WF01D).[118] Patients will also incur the costs of the imaging test received:

- Colonoscopy: The model assumes to include the costs of a colonoscopy, based on the weighted average costs of colonoscopy with biopsy, therapeutic colonoscopy and diagnostic colonoscopy (codes FE30Z, FE31Z and FE32Z) from NHS Reference costs,[118] plus the costs of same day attendance imaging admission or attendance (code RD97Z) and a follow-up appointment with

gastroenterologist (based on weighted average cost for face-to-face and non-face-to-face attendances with a gastroenterologist - codes 301, WF01A and WF01C). The total cost of a colonoscopy was estimated to be £1,003.34.

- CTC: The model includes the costs of a CTC scan (code RD61Z) in addition to the costs of a same day attendance imaging admission or attendance (code RD97Z) and a follow-up appointment (based on weighted average cost for face-to face and non-face-to-face attendances with a gastroenterologist - codes 301, WF01A and WF01C).[118] The total cost of a CTC was estimated to be £341.17.
- Other non-invasive investigations: The model includes the cost of a CT (based on the weighted average cost of CT scans of one or more areas with and without contrast available for adults - codes RD20A to RD27Z)[118] for 80% of patients (assumption made based on clinical opinion from EAG's advisors – see Appendix 11), and one additional appointment with a consultant gastroenterologist to discuss treatment action for all patients (weighted average of face-to face and non-face-to-face attendances with a gastroenterologist, follow-up codes R301, WF01A and WF01C). The total cost of other non-invasive investigations was estimated to be £256.29 per patient.

Costs of Watch and Wait

Patients who receive a FIT result lower than t or t_{high} or who do not complete the test are followed-up in primary care under the 'watch and wait' pathway. These patients are assumed to receive an additional 1.9 GP appointments, based on estimated data from COLOFIT modelling team based on Lyratzopoulos *et al.* (2013),[119] costing £68.40.[121] Additional costs during watchful waiting was estimated by calculating the weighted mean of potential costs.

Based on NCRAS data on routes to diagnosis for England 2018,[120] 22% of patients with underlying bowel disease (which was estimated to be 7.8% of patients presenting to the GP with symptoms of CRC) are assumed to incur on the costs of A&E attendance (£296.88) based on the weighted average of all A&E attendances excluding dental and dead on arrival – codes VB01Z to VB11Z from NHS Reference Costs).[118] The remaining 78% of patients with underlying bowel disease are assumed to be eventually detected and referred by the GP to receive an initial appointment with a consultant gastroenterologist (£186.48) and a colonoscopy (£1003.34).[118] Patients with NSBP are assumed not to re-present to their GP and not to incur costs further to the additional 1.9 GP appointments.

The weighted mean cost per on the 'watch and wait' pathway was estimated to be £145.97.

Costs associated with 'Repeat FIT'

Patients invited to complete a second FIT (under the 'repeat FIT' pathway) are assumed to incur the costs of the test, including the costs of the additional samples needed for those who complete it. The proportion of completed FITs that need resampling was assumed to be the same as for the first test. The total cost of repeat FIT per patient also includes 1.9 additional GP appointments (unit cost taken from PSSRU 2022).[121] Patients who do not complete the second FIT or who receive a 'negative' result are assumed to incur in the costs of 'Watch and Wait', whilst patients who complete it and receive a 'positive' result (based on the accuracy estimates of the test for CRC) are assumed to incur in the cost of a colonoscopy and an appointment with a consultant gastroenterologist (both based on NHS Reference Costs 2021/22 using the same codes listed in the costs of referrals).[118] The cost per patient of the 'repeat FIT' pathway varies depending on the FIT brand received and threshold used, and is estimated to be, for example, between £339.92 and £586.21 for the threshold 10 µg/g (within a same brand, the cost decreases with higher thresholds).

Costs of treating complications related to colonoscopy

The costs of treating complications related to colonoscopy are summarised in Table 49.

Table 49: Costs related to complications of colonoscopy

Complication	Unit costs	Source
Bleeding	£1,695.45	NHS Reference Costs 2021/22, weighted average cost of all Gastrointestinal Bleed procedures with multiple, single or no interventions (Codes FD03A to FD03H)[118]
Perforation	£6,299.74	NHS Reference Costs 2021/22, weighted average cost of all major large intestine procedures in adults (19+, codes FF34A to FF34C)[118]
Death after perforation	£0.00	Assumption that the costs of perforation already capture the costs incurred before patient's death

Costs associated with treating the underlying conditions (lifetime costs for CRC, IBD and AAs)

Appendix 12 presents details on the costs' estimates used in the Whyte *et al.* (2023) model for CRC and AA patients.

For IBD patients, the lifetime costs were estimated from annual costs related to the disease treatment based the proportion of patients by disease severity and type of disease reported in Table ***47*** and unit costs from Ghosh et al. (2015).[107] The annual cost of treatment for UC was estimated to be £3,083.94 and for CD £6,156.44. Based on the relative incidence of UC versus CD from Pasvol *et al*. (2020)[106] of 0.6, the annual cost of IBD was estimated to be £4,297.70, which was uplifted to 2022 prices using the NHS cost inflation index (NHSCII)[117] to £ £5,015.75. The impact of delay on diagnosis of IBD was estimated to be £399.66, based on the difference in costs between severe relapse and milder forms of UC and CD, applied for 2 years since diagnosis, to those patients who are diagnosed within the 'watch

and wait', 'repeat FIT' or are missed by the diagnostic tests. Patients with NSBP are assumed not to incur in any long-term costs.

5.3.5. Methods for model evaluation

The health outcomes and costs of each testing strategy were generated for each brand of FIT based on each threshold and pair of thresholds being evaluated (with exception of Intervention 3, where the threshold of 10µg/g currently in place under DG30 was used). The total outcomes were evaluated against each other in full incremental analyses. The cost-effectiveness of each test brand was also compared against each other for selected thresholds. Results based on net monetary benefit (NMB) were also generated. Central estimates of cost-effectiveness were based on the expectation of the mean. Uncertainty was evaluated using probabilistic sensitivity analysis (PSA) and deterministic sensitivity analyses (DSA). PSA was undertaken using simple Monte Carlo sampling methods (1,000 samples). The choice of distribution assumed for each group of parameters in the model is summarised in Table 50.

Table 50: Distributions used in EAG probabilistic analyses

Model parameter group	Model parameter	Distribution	EAG comments
Patient characteristics	age	Fixed	
	Proportion of female	Beta	
Settings	Discount rates (QALYs and costs)	fixed	
Disease prevalence (overall population and high- and low-risk groups)	CRC prevalence		CODA sampling, for overall population an estimate based on high and low-risk groups was estimated. Samples for low-risk based on beta distribution.
	AAs prevalence		
	IBD prevalence		
	Proportion of high risk-patients	beta	
Disease stage/severity distribution	CRC stage distribution	See Appendix 12	
	UC/CD severity	fixed	assumption
Tests' accuracy	FIT	CODA samples / beta	CODA samples when point estimates from EAG's clinical review and analysis; beta when data from unique study
	COL	Beta/fixed	Sensitivities were samples, whilst specificity were assumed to be 1.0
	CTC	beta	
	Other non-invasive interventions	fixed	Assumed to be 1.0
Safety netting and 'intermediate group' pathways	Probability of receiving each of the pathways following a FIT result <t or $<t_{low}$	fixed	
	Proportion of patients receiving each of the pathways following a FIT result $>t_{low}$ and $<t_{high}$	fixed	
Interventions received in 2WW and	Proportion of patients of receiving each imaging test in 2WW and 18WW referrals	fixed	
	Complications after colonoscopies	beta	

18WW referrals			
Time to diagnosis and length of delay	Time to diagnosis	fixed	Varied in scenario analysis
	Proportion of change in total referrals (2WW+18WW) in comparison to intervention 3 (applied to time to diagnosis in interventions 1 and 2)	normal	SD assumed to be 0.1
Costs	FIT total costs (completed tests)	normal	SD assumed to be 0.1
	FIT costs (non- completed tests)	fixed	
	Repeat FIT total costs (completed tests)	normal	SD assumed to be 0.1
	Repeat FIT costs (non- completed tests)	fixed	
	FIT return rate	beta	
	Referral - initial appointment	normal	SD assumed to be 0.1
	Colonoscopy	normal	SD assumed to be 0.1
	CTC	normal	SD assumed to be 0.1
	Other non-invasive interventions	normal	SD assumed to be 0.1
	Watch and wait	normal	SD assumed to be 0.1
	COL complications	normal	SD assumed to be 0.1, exception for death after perforation which is assumed fixed (£0.0)
	Annual cost treatment IBD	normal	
	Increased treatment cost for IBD due to delay in diagnosis	normal	
Long term STM model	Lifetime survival, QALYs and costs (CRC and AAs)	See Appendix 12	
	IBD and NSBP survival	fixed	Based on general population's Lifetables
HRQoL	General population utility values (by age and sex)	fixed	
	IBD utility value	beta	
	Utility multiplier for IBD delayed diagnosis	normal	
	Utility loss due to COL complications	normal	

AAs – advanced adenomas; CRC – colorectal cancer; FIT - Faecal Immunochemical Test; IBD – inflammatory bowel disease; SD – standard deviation; STM – state transition model

5.3.5.1. Scenario analyses

The EAG undertook a number of deterministic scenario analyses, in order to test the robustness of the results generated by the model to changes in key assumptions. These included:

- Testing the influence of the estimates for time to diagnosis on the results: the EAG tested two different scenarios, where the lengths of time to diagnosis for each possible pathway in the model were assumed to be shorter (scenario 1) and longer (scenario 2) than in the model's base-case. The estimates were based on clinical opinion from the EAG's advisors (see Table 45).
- Inclusion of a QALY loss equivalent of one day of full health for the same age and sex-match, to estimate the patient's anxiety and inconvenience associated to receiving a colonoscopy. This QALY loss was applied to all patients who received a colonoscopy in the model.
- Inclusion of a QALY loss equivalent of one day of full health for each month of diagnostic delay. This QALY loss was applied to all patients with underlying disease (CRC, IBD, AAs) and also to those without underlying disease who have been referred to 2WW and 18WW.

- Testing the use of DUAL FIT instead of a single FIT. This scenario assumed to have accuracy data for the DUAL FIT taken from the EAG's clinical review for HM JACKarc (see Section 4.3.7) and the unit costs of the first FIT for Intervention 1 to be the double for that brand.
- Removing IBD and AAs from the model: given the uncertainty around the parameters for the other bowel pathologies included in the model, the EAG tested the removal of IBD and AA, by assuming they have zero prevalence.
- Lower return rate for FIT: the EAG, given the value for the return rate of FIT in the base-case is 0.91 and it might be considered high in the primary care context, the EAG tested using a second source from Moss *et al.*(2017)[125] of 0.664.
- Alternative assumption about diagnostic accuracy of FIT in the DG30 low-risk patients in Intervention 3 based on the EAG's systematic review. The values for sensitivity and specificity for DGD30 low-risk patients in Intervention 3 in this scenario is presented in
- Table ***51***.

Table 51: Accuracy estimates used in Scenario 8 for detection of CRC

	Sensitivity	95%CI	Specificity	95%CI
CRC	0.910	0.815-0.978	0.911	0.769-0.983

- Alterative assumption of increased resource use in terms of GP appointments for patients with NSBP undertaking watch and wait and Repeat FIT pathways. In this alternative scenario, the model assumes that patients without underlying disease that are not directly referred would receive one additional GP appointment.
- Alternative assumption for the cost of the test at current recommendations: in this scenario, the EAG changed the unit cost of FIT for Intervention 3, from the lowest cost available to a weighted average of the unit costs informed by the manufacturers, using the number of studies that were included in the statistical analyses used to inform the accuracy of the test for this intervention (unit cost changed from £2.31 to £3.28).
- The EAG also ran a scenario where the FIT was assumed to have perfect accuracy (sensitivity and specificity = 1.0) and where all patients return the test (return rate =1) to test an extreme scenario where no patients are missed by test or wrongly sent to 2WW.

5.3.6. Model verification and validation

The EAG undertook a number of measures to ensure the validity of the model.

- Peer review of the economic analysis by a modeller not involved in the assessment
- Verification and scrutiny of the executable model by two model developers
- Double-checking of the accuracy of all model inputs against sources

- Comparison of model results using point estimates of parameters and the expectation of the mean
- Comparison of mean of all probabilistic parameter samples against point estimates of parameters
- Examination of all identified sources of discrepancy
- Model testing using sensitivity analysis and use of extreme parameter values.

5.3.7. Cost-effectiveness results

Four key sets of results have been produced which include high or low safety netting intensity (see Section 5.3.4.4) and assuming a willingness to pay threshold of £20,000 or £30,000 per QALY gained.[94] To allow multiple results to be shown on figures an incremental net monetary benefit approach (iNMB) has been used by the EAG which requires specification of the assumed threshold. iNMB is defined as the cost per QALY gained threshold multiplied by the incremental QALY gain minus the incremental cost;[126] under this framework the largest estimated iNMB is deemed to be the most cost-effective strategy, which could be zero if the benchmark intervention is most cost-effective. The absolute loss (valued in terms of cost) of moving to a different strategy calculated by comparing the estimated iNMBs.

NMB also has an advantage that if the assumed costs are believed to be imprecise then the level of additional or reduced costs (for example, the additional costs of GP appointments incurred over the base case) can be directly applied to the NMB values. It is for this reason that NMB is preferred to net health benefits although the conclusions are identical whichever metric is used. The NMB values presented are per person.

The conclusions from all four analyses are similar and therefore only one set of results are presented in the main text, with the results for the remaining three analyses contained in Appendix 14. The chosen combination uses a low safety netting approach and a threshold of £20,000 per QALY gained as this is most different from current standard of care and the lower threshold as there is likely considerable uncertainty in the ICER given the small differences in QALYs between strategies.

The structure for presenting (and interpreting) results is as follows:

1) A figure depicting the iNMBs for each of the seven tests at selected thresholds when only one threshold is used (denoted Intervention 1).
2) A figure depicting the iNMBs for the five tests with sufficient data at selected thresholds when two thresholds are used t_{low} and t_{high} (denoted Intervention 2).
3) Tables for each test that display summarised data relating to the clinical and cost-effectiveness of each test compared with current care (denoted Intervention 3). The EAG has

selected the data that it considers to be most pertinent for decision making; other data can be provided by the EAG on request.

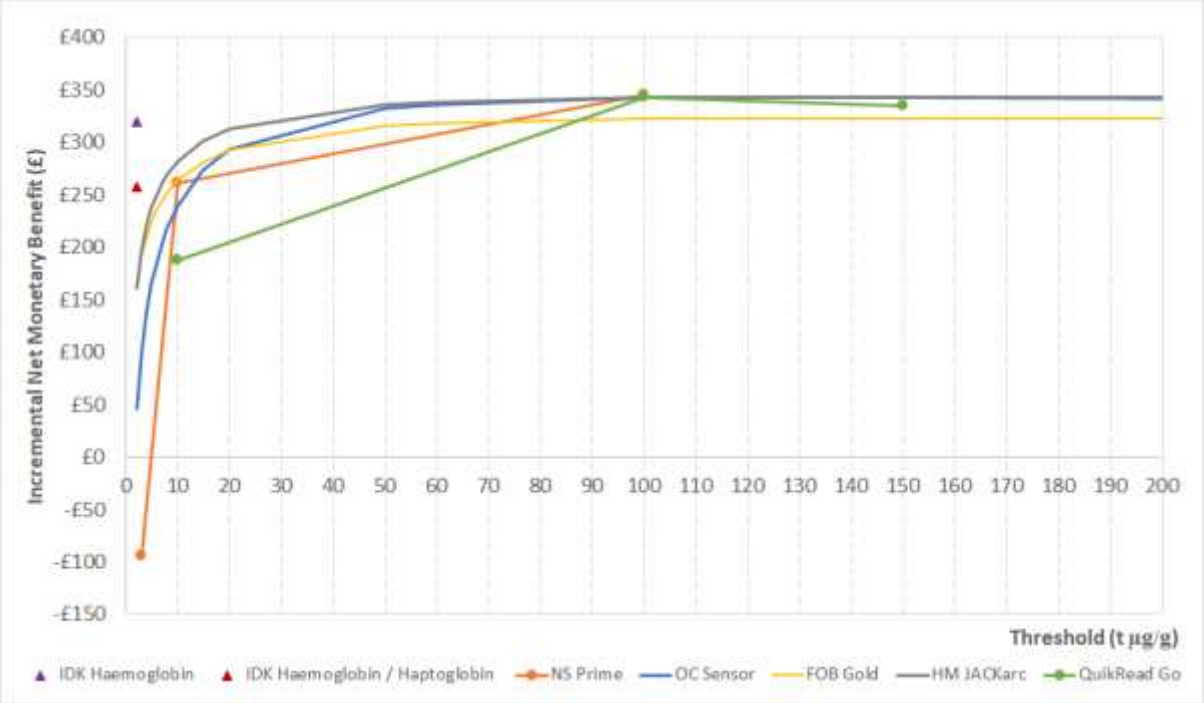
The EAG ran PSA for selected tests and thresholds; these indicated that the NMB values differed by less than £10 on average. Given the linearity of the results and the timescales of the project the EAG has deemed that presenting only deterministic results would not influence decision making.

5.3.7.1. Results assuming a threshold of £20,000 per QALY gained and low intensity safety netting threshold

5.3.7.2. The iNMBs of the seven tests using one threshold

Figure ***18*** shows the iNMB for the diagnostic strategies for FIT when using one threshold (Intervention 1). The iNMBs for higher threshold values to be in the region of £300 to £350 for all tests. With the exception of NS Prime at a low threshold all tests have positive iNMBs compared with current practice. The reason for the negative iNMB for NS Prime at a threshold of 3µg/g is due to the poor estimated specificity of the test at this threshold (0.319) which came from one study Benton *et al*.[46] which had a very small number of events. This results in a large number of patients being referred to colonoscopy. iNMB values in Figure 18 have been interpolated resulting in straight lines where there is a distance between thresholds (for example, 10µg/g and 100µg/g). Whilst the highest iNMB values appear to be around a threshold of 100µg/g, the iNMB loss using a threshold of 50µg/g is slight. Whilst for the majority of tests there is a noticeable reduction in iNMB at a threshold of 10µg/g, the results show that all tests used at this threshold have a higher iNMB than current practice. Given the uncertainty in the model input parameters the EAG notes that the generated comparisons between thresholds for a particular test, or between tests themselves, may not be robust although broad conclusions are likely robust.

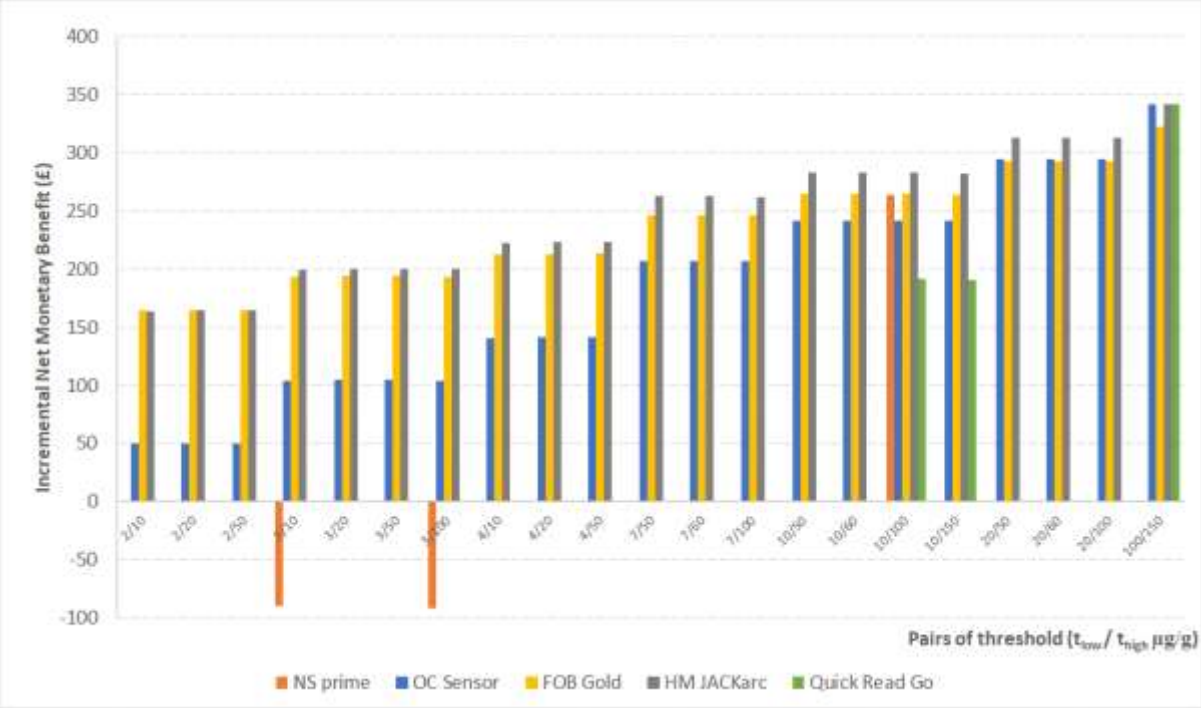
Figure 18: NMB for Intervention 1 assuming a threshold of £20,000 per QALY gained and low intensity safety netting


£400
£350
£300
£250
£200
£150
£100
£50
£0
-£50
-£100
-£150
Incremental Net Monetary Benefit (£)
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200
Threshold (t µg/g)
IDK Haemoglobin
IDK Haemoglobin / Haptoglobin
NS Prime
OC Sensor
FOB Gold
HM JACKarc
QuikRead Go

5.3.7.3. The iNMBs of the five tests using two thresholds

Figure 19 shows the iNMB when using FIT strategies with two thresholds. The iNMBs for higher threshold values appears to be in the region of £300 to £350 for all tests. These iNMBs are lower than when one threshold is used (Figure 18). All iNMBs are positive except for when the value for t_{low} is set to 3µg/g for NS Prime, which is due to the low specificity for this test at this threshold. Many pairs of combinations have reasonably similar iNMB values given the underlying uncertainty, for instance using paired values of 7µg/g and 50µg/g compared with using 10µg/g and 100µg/g.

Figure 19: NMB for Intervention 2 assuming a threshold of £20,000 per QALY gained and low intensity safety netting



5.3.7.4. Tabulated results for each test

Table 52 to ■Southwest quadrant ICER. - AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table *61* show tabulated results for each test (at selected thresholds). The tests are presented in alphabetical order. The vast majority of ICERs presented in these tables lie in the South-West quadrant of the cost-effectiveness plane, which is due to the tests generating marginally fewer QALYs and lower costs than current practice. The explanation for this phenomenon, using FOB Gold at a threshold of 10µg/g as an example (as it is first alphabetically) follows.

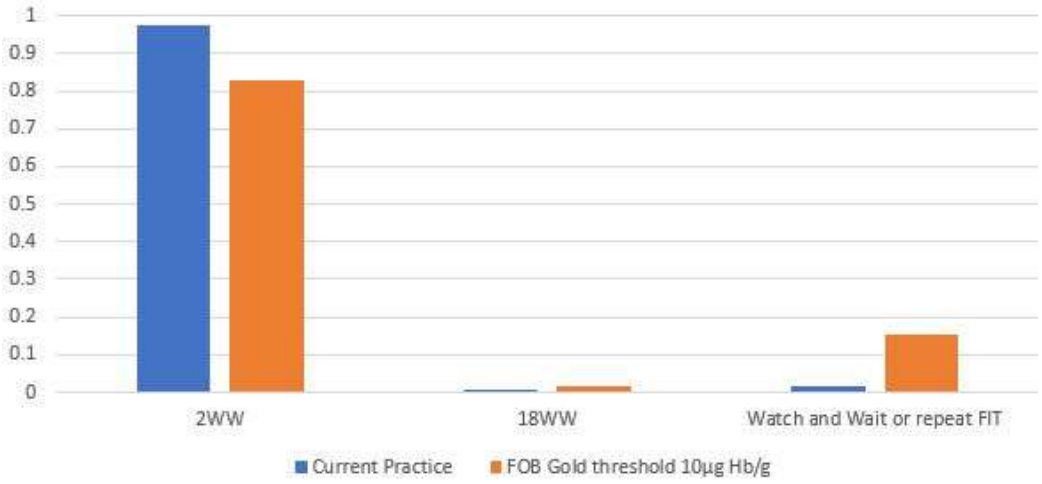
Illustrated example of why QALYs are marginally decreased using FIT

Under current practice, a large number of patients are referred to receive colonoscopy (0.623 colonoscopies per patient) (Table 52). For FOB Gold at a threshold of 10µg/g, some patients with bowel disease will be missed due to the comparatively lower sensitivity of the test and fewer patients will receive colonoscopy, with an estimated 0.307 colonoscopies per patient (Table 52). Undertaking more colonoscopies increases the probability that patients with bowel disease will be detected and receive appropriate treatment. Patients with underlying CRC who do not receive colonoscopy experience delays which result in worse outcomes and lower QALYs for the cohort.

Whilst there is a benefit in quicker time to diagnosis for those in the 2WW and 18WW pathways using FIT due to less demand on colonoscopy resources this is not sufficient to outweigh the losses associated with significantly delayed diagnosis. The mean time to a diagnosis of CRC is 1.384 months in current practice and 2.668 months for FOB Gold at a threshold of 10µg/g (Table 52). The later average diagnosis of CRC (and similarly for AAs and IBD) mean that QALYs for the cohort are decreased from 10.895 in current practice to 10.893 for FOB Gold at a threshold of 10µg/g (Table 52), which is less than 1 day of full health for all patients in the cohort.

The different proportions of patients with CRC diagnosed in the 2WW pathway, the 18WW pathway and the watch and wait and repeat FIT combined by current practice and FOB Gold at a threshold of 10µg/g are shown in Figure 20.

Figure 20: The proportion of patients diagnosed by category



Illustrated example of why costs are lower when using FIT

Whilst there will be an increase in costs associated with the use of FIT there is a consequential decrease in the number of colonoscopies undertaken. As noted, in the section above, the estimated average number of colonoscopies undertaken per person was 0.623 for current practice and 0.307 for FOB Gold at a threshold of 10µg/g. This reduction in colonoscopy usage generates a considerable saving, which drives an overall reduction of costs from £3142 in current care to £2836 for FOB Gold at a threshold of 10µg/g (Table 52).

Combining the estimated implications for QALYs accrued and costs incurred, the ICER for FOB Gold at a threshold of 10µg/g is calculated to be £142,533 (Table 52), although this is in the South West quadrant, indicating that for every QALY yielded there would be a saving of £142,533. Alternatively, this could be viewed as current care having an ICER of £142,533 compared with FOB Gold at a threshold of 10µg/g, which is greater than the thresholds of £20,000 or £30,000 published by NICE.[94]

Table 52: Tabulated results for FOB Gold using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.167	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.894	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2954	2899	2883	2859	2836	2795	2751	2743	2723	3142
ICER (pairwise, vs Intervention 3)▫ (£)	145,571	154,995	154,119	149,918	142,533	126,587	102,807	98,287	86,628	-
NMB λ=20,000 (vs Int 3) (£)	162	211	225	245	263	292	315	318	322	-
NMB λ=30,000 (vs Int 3) (£)	149	195	209	226	242	265	277	277	274	-
Number of 2WW referrals (total)	0.385	0.325	0.307	0.282	0.256	0.213	0.168	0.160	0.141	0.644
Number of 18WW referrals (total)	0.065	0.071	0.073	0.076	0.078	0.083	0.088	0.088	0.090	0.037
Number of Repeat FITs (total)	0.065	0.071	0.073	0.076	0.078	0.083	0.088	0.088	0.090	0.037
Number of Watch and Wait (total) (total)	0.485	0.533	0.547	0.567	0.587	0.621	0.657	0.663	0.678	0.281
Number of COLs (total)	0.413	0.364	0.349	0.328	0.307	0.272	0.235	0.229	0.212	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.9%	41.8%	44.2%	47.6%	50.9%	56.5%	62.5%	63.5%	66.1%	-
Reduction in number of referrals (2WW only)	40.2%	49.5%	52.3%	56.2%	60.2%	66.9%	73.9%	75.1%	78.1%	-
Increase in number of referrals (18WW only)▫▫	72.6%	89.5%	94.5%	101.7%	108.8%	120.9%	133.6%	135.8%	141.3%	-
Reduction in number of COLs	33.8%	41.7%	44.0%	47.4%	50.7%	56.3%	62.3%	63.3%	65.9%	-
Mean time to diagnosis - CRC	2.204	2.339	2.400	2.513	2.668	3.087	3.950	4.169	4.865	1.384
Mean time to diagnosis - AAs	4.453	5.133	5.398	5.833	6.355	7.224	8.346	8.563	9.117	1.956
Mean time to diagnosis - IBD	2.944	3.361	3.521	3.788	4.106	4.814	5.837	6.044	6.607	2.044

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 53: Tabulated results for FOB Gold using two thresholds

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.167	14.167	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.165	14.165	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.892	10.892	10.895
Costs (£)	2951	2949	2948	2897	2896	2894	2855	2854	2832	2831	2793	2792	3142
ICER (pairwise, vs Intervention 3)□ (£)	141,851	138,599	132,652	152,733	150,052	144,984	143,209	139,404	137,816	134,686	124,564	122,479	-
NMB λ=20,000 (vs Int 3) (£)	164	165	165	213	213	213	246	246	264	264	293	293	-
NMB λ=30,000 (vs Int 3) (£)	151	151	150	197	197	196	226	226	242	241	265	264	-
Number of 2WW referrals (total)	0.365	0.358	0.351	0.314	0.308	0.300	0.264	0.259	0.242	0.238	0.206	0.202	0.644
Number of 18WW referrals (total)	0.078	0.083	0.088	0.078	0.083	0.088	0.088	0.090	0.088	0.090	0.088	0.090	0.037
Number of Repeat FITs (total)	0.071	0.074	0.076	0.075	0.077	0.079	0.082	0.083	0.083	0.084	0.085	0.087	0.037
Number of Watch and Wait (total) (total)	0.485	0.485	0.485	0.533	0.533	0.533	0.567	0.567	0.587	0.587	0.621	0.621	0.281
Number of COLs (total)	0.406	0.404	0.402	0.360	0.358	0.356	0.323	0.321	0.303	0.302	0.270	0.269	0.623
Reduction in number of referrals (total - 2WW + 18WW)	34.9%	35.3%	35.6%	42.4%	42.7%	43.1%	48.4%	48.6%	51.6%	51.8%	56.9%	57.1%	-
Reduction in number of referrals (2WW only)	43.3%	44.4%	45.5%	51.2%	52.2%	53.3%	59.0%	59.7%	62.4%	63.0%	68.0%	68.6%	-
Increase in number of referrals (18WW only)	108.8%	120.9%	133.6%	108.8%	120.9%	133.6%	133.6%	141.3%	133.6%	141.3%	133.6%	141.3%	-
Increase in number of repeat FITs	90.7%	96.7%	103.1%	99.2%	105.2%	111.6%	117.7%	121.5%	121.2%	125.1%	127.3%	131.1%	-
Increase in number of watch and waits	72.6%	72.6%	2.6%	89.5%	89.5%	89.5%	101.7%	101.7%	108.8%	108.8%	120.9%	120.9%	-
Reduction in number of COLs	34.8%	35.1%	35.5%	42.2%	42.5%	42.9%	48.3%	48.5%	51.4%	51.6%	56.7%	56.9%	-
Mean time to diagnosis - CRC	2.236	2.265	2.323	2.360	2.386	2.440	2.599	2.654	2.742	2.795	3.135	3.184	1.384
Mean time to diagnosis - AAs	4.580	4.637	4.711	5.209	5.263	5.332	5.980	6.024	6.467	6.511	7.283	7.324	1.956
Mean time to diagnosis - IBD	3.023	3.071	3.141	3.408	3.453	3.518	3.911	3.958	4.207	4.252	4.871	4.913	2.044

▪*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 54: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold (µg/g)									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2953	2888	2868	2841	2815	2772	2728	2,721	2703	3142
ICER (pairwise, vs Intervention 3)▪ (£)	136,236	152,176	152,565	149,960	143,701	129,021	106,525	102,213	91,140	-
NMB λ=20,000 (vs Int 3) (£)	161	221	238	260	281	313	336	338	342	-
NMB λ=30,000 (vs Int 3) (£)	147	204	220	240	259	284	297	297	294	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)▪▪	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

▪Southwest quadrant ICER; ▪▪ Also the value for increased repeat FITs and increased number of watch and waits
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 55: Tabulated results for HM JACKarc using two thresholds

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.165	14.165	14.165	14.165	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.892	10.892	10.892	10.892	10.895
Costs (£)	2949	2948	2946	2885	2884	2882	2837	2836	2812	2811	2770	2769	3142
ICER (pairwise, vs Intervention 3) (£)	133,360	130,590	125,636	150,262	147,898	143,524	144,094	140,891	139,564	136,918	127,237	125,455	-
NMB λ=20,000 (vs Int 3) (£)	164	164	164	223	223	223	262	262	283	283	313	313	-
NMB λ=30,000 (vs Int 3) (£)	149	150	149	205	206	205	241	241	259	259	284	284	-
Number of 2WW referrals (total)	0.363	0.355	0.348	0.302	0.295	0.288	0.246	0.242	0.222	0.219	0.184	0.180	0.644
Number of 18WW referrals (total)	0.080	0.085	0.090	0.080	0.085	0.090	0.090	0.092	0.090	0.092	0.090	0.092	0.037
Number of Repeat FITs (total)	0.072	0.075	0.077	0.076	0.079	0.081	0.084	0.085	0.085	0.086	0.087	0.089	0.037
Number of Watch and Wait (total) (total)	0.485	0.485	0.485	0.541	0.541	0.541	0.581	0.581	0.603	0.603	0.638	0.638	0.281
Number of COLs (total)	0.406	0.404	0.402	0.351	0.349	0.347	0.308	0.307	0.287	0.286	0.252	0.251	0.623
Reduction in number of referrals (total - 2WW + 18WW)	35.0%	35.3%	35.7%	43.9%	44.2%	44.6%	50.7%	50.9%	54.2%	54.4%	59.8%	60.0%	-
Reduction in number of referrals (2WW only)	43.7%	44.8%	45.9%	53.1%	54.2%	55.3%	61.8%	62.4%	65.5%	66.1%	71.4%	72.0%	-
Increase in number of referrals (18WW only)	114.4%	127.1%	139.6%	114.4%	127.1%	139.6%	139.6%	146.5%	139.6%	146.5%	139.6%	146.5%	-
Increase in number of repeat FITs	93.4%	99.7%	106.0%	103.4%	109.8%	116.1%	123.1%	126.5%	127.0%	130.5%	133.4%	136.8%	-
Increase in number of watch and waits	72.4%	72.4%	72.4%	92.5%	92.5%	92.5%	106.5%	106.5%	114.4%	114.4%	127.1%	127.1%	-
Reduction in number of COLs	34.8%	35.2%	35.5%	43.7%	44.0%	44.4%	50.5%	50.7%	54.0%	54.2%	59.6%	59.8%	-
Mean time to diagnosis - CRC	2.342	2.371	2.426	2.482	2.508	2.558	2.732	2.779	2.883	2.928	3.279	3.321	1.384
Mean time to diagnosis - AAs	4.579	4.637	4.711	5.200	5.253	5.321	5.963	6.007	6.449	6.491	7.264	7.303	1.956

Mean time to diagnosis - IBD	3.023	3.071	3.140	3.400	3.444	3.508	3.896	3.941	4.189	4.233	4.852	4.893	2.044

▪*Southwest quadrant ICER.*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 56: Tabulated results for IDK Haemoglobin using one threshold

	Int 1: FIT 1 threshold	Int 3: DG30& NG12
t (µg/g)	**2**	**10**
LYs	14.165	14.168
QALYs	10.893	10.895
Costs (£)	2783	3142
ICER (pairwise, vs Intervention 3)▪ (£)	182,163	-
NMB λ=20,000 (vs Int 3) (£)	319	-
NMB λ=30,000 (vs Int 3) (£)	300	-
Number of 2WW referrals (total)	0.201	0.644
Number of 18WW referrals (total)	0.084	0.037
Number of Repeat FITs (total)	0.084	0.037
Number of Watch and Wait (total) (total)	0.631	0.281
Number of COLs (total)	0.263	0.623
Reduction in number of referrals (total - 2WW + 18WW)	58.1%	-
Reduction in number of referrals (2WW only)	68.8%	-
Increase in number of referrals (18WW only) ▪▪	124.4%	-
Reduction in number of COLs	57.9%	-
Mean time to diagnosis - CRC	3.027	1.384
Mean time to diagnosis - AAs	4.328	1.956
Mean time to diagnosis - IBD	2.814	2.044

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 57: Tabulated results for IDK Haemoglobin using one threshold

	Int 1: FIT 1threshold	Int 3: DG30& NG12
t (µg/g)	**2**	**10**
LYs	14.164	14.168
QALYs	10.892	10.895
Costs (£)	2836	3142
ICER (pairwise, vs Intervention 3)▪ (£)	126,916	-
NMB λ=20,000 (vs Int 3) (£)	257	-
NMB λ=30,000 (vs Int 3) (£)	233	-
Number of 2WW referrals (total)	0.258	0.644
Number of 18WW referrals (total)	0.078	0.037
Number of Repeat FITs (total)	0.078	0.037
Number of Watch and Wait (total) (total)	0.585	0.281
Number of COLs (total)	0.309	0.623
Reduction in number of referrals (total - 2WW + 18WW)	0.506	-
Reduction in number of referrals (2WW only)	59.9%	-
Increase in number of referrals (18WW only) ▪▪	108.3%	-
Reduction in number of COLs	50.4%	-
Mean time to diagnosis - CRC	3.477	1.384
Mean time to diagnosis - AAs	4.367	1.956
Mean time to diagnosis - IBD	2.855	2.044

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 58: Tabulated results for NS Prime

	Int 1: FIT 1 threshold			Int 2: FIT 2 thresholds			Int 3: DG30&NG12 (FIT T=10)
Threshold - t or t_{low}/t_{high} (µg/g)	**3**	**10**	**100**	**3/10**	**3/100**	**10/100**	**10**
LYs	14.164	14.162	14.160	14.164	14.164	14.162	14.168
QALYs	10.892	10.891	10.889	10.892	10.892	10.891	10.895
Costs (£)	3183	2804	2684	3177	3175	2800	3142
ICER (pairwise, vs Intervention 3) (£)	Dominated	88,041▫	81,074▫	Dominated	Dominated	86,949▫	-
NMB λ=20,000 (vs Int 3) (£)	-94	261	345	-90	-92	264	-
NMB λ=30,000 (vs Int 3) (£)	-120	223	288	-118	-121	224	-
Number of 2WW referrals (total)	0.645	0.226	0.101	0.579	0.559	0.206	0.644
Number of 18WW referrals (total)	0.037	0.081	0.095	0.081	0.095	0.095	0.037
Number of Repeat FITs (total)	0.037	0.081	0.095	0.059	0.066	0.088	0.037
Number of Watch and Wait (total) (total)	0.280	0.611	0.710	0.280	0.280	0.611	0.281
Number of COLs (total)	0.624	0.282	0.180	0.604	0.598	0.276	0.623
Reduction in number of referrals (total - 2WW + 18WW)	-0.2%	54.9%	71.3%	3.1%	4.0%	55.9%	-
Reduction in number of referrals (2WW)	-0.2%	64.9%	84.3%	10.1%	13.1%	68.0%	-
Increase in number of referrals (18WW)	0.4%	-117.4%	-152.5%	117.4%	152.5%	152.5%	-
Increase in number of repeat FITs	0.4%	-117.4%	-152.5%	58.5%	76.0%	135.0%	-
Increase in number of watch and waits	0.4%	-117.4%	-152.5%	-0.4%	-0.4%	117.4%	-
Reduction in number of COLs	-0.1%	54.7%	71.1%	3.2%	4.1%	55.7%	-
Mean time to diagnosis - CRC	3.451	4.513	5.762	3.550	3.660	4.582	1.384
Mean time to diagnosis - AAs	5.035	6.331	9.081	5.154	5.391	6.480	1.956
Mean time to diagnosis - IBD	3.385	4.082	6.572	3.454	3.671	4.222	2.044

▪*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 59: Tabulated results for OC-Sensor using one threshold

	Int 1: FIT 1 threshold (µg/g)									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.161	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	£3,066	£2,970	£2,940	£2,898	£2,857	£2,791	£2,731	£2,722	£2,701	3142
ICER (pairwise, vs Intervention 3)□ (£)	51,242	99,829	110,117	120,293	124,509	121,539	103,972	99,908	89,062	-
NMB λ=20,000 (vs Int 3) (£)	£46	£137	£165	£204	£239	£293	£332	£336	£342	-
NMB λ=30,000 (vs Int 3) (£)	£31	£120	£147	£183	£217	£264	£292	£294	£292	-
Number of 2WW referrals (total)	0.510	0.402	0.369	0.322	0.278	0.209	0.147	0.138	0.117	0.644
Number of 18WW referrals (total)	0.052	0.063	0.066	0.071	0.076	0.083	0.090	0.091	0.093	0.037
Number of Repeat FITs (total)	0.052	0.063	0.066	0.071	0.076	0.083	0.090	0.091	0.093	0.037
Number of Watch and Wait (total) (total)	0.387	0.472	0.498	0.535	0.570	0.625	0.673	0.680	0.697	0.281
Number of COLs (total)	0.514	0.426	0.399	0.361	0.325	0.269	0.218	0.211	0.193	0.623
Reduction in number of referrals (total - 2WW + 18WW)	17.6%	31.8%	36.1%	42.2%	48.0%	57.1%	65.2%	66.4%	69.1%	-
Reduction in number of referrals (2WW only)	20.8%	37.6%	42.8%	50.0%	56.8%	67.6%	77.1%	78.5%	81.8%	-
Increase in number of referrals (18WW only)□□	37.7%	68.1%	77.3%	90.3%	102.7%	122.2%	139.5%	142.0%	147.9%	-
Reduction in number of COLs	17.6%	31.7%	36.0%	42.1%	47.9%	56.9%	65.0%	66.2%	69.0%	-
Mean time to diagnosis - CRC	2.347	2.475	2.536	2.653	2.812	3.247	4.118	4.335	5.004	1.384
Mean time to diagnosis - AAs	4.537	5.187	5.442	5.863	6.372	7.220	8.328	8.544	9.096	1.956
Mean time to diagnosis - IBD	3.031	3.416	3.565	3.818	4.122	4.811	5.820	6.025	6.586	2.044

▪Southwest quadrant ICER; □□ Also the value for increased repeat FITs and increased number of watch and waits

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 60: Tabulated results for OC-Sensor using two thresholds

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.165	14.165	14.165	14.165	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.892	10.892	10.892	10.892	10.895
Costs (£)	3061	3060	3059	2966	2965	2963	2892	2891	2852	2851	2789	2788	3142
ICER (pairwise, vs Intervention 3)□ (£)	52,217	51,617	49,889	99,463	98,233	95,306	116,119	113,275	121,240	118,728	119,955	118,170	-
NMB λ=20,000 (vs Int 3) (£)	£50	£50	£50	£140	£141	£141	£207	£207	£242	£242	£294	£294	-
NMB λ=30,000 (vs Int 3) (£)	£34	£34	£33	£122	£123	£123	£185	£185	£218	£217	£265	£264	-
Number of 2WW referrals (total)	0.473	0.462	0.453	0.382	0.371	0.361	0.295	0.290	0.257	0.253	0.199	0.194	0.644
Number of 18WW referrals (total)	0.076	0.083	0.090	0.076	0.083	0.090	0.090	0.093	0.090	0.093	0.090	0.093	0.037
Number of Repeat FITs (total)	0.064	0.067	0.071	0.069	0.073	0.076	0.081	0.082	0.083	0.084	0.087	0.088	0.037
Number of Watch and Wait (total) (total)	0.387	0.387	0.387	0.472	0.472	0.472	0.535	0.535	0.570	0.570	0.625	0.625	0.281
Number of COLs (total)	0.503	0.499	0.496	0.420	0.416	0.413	0.353	0.351	0.319	0.317	0.266	0.264	0.623
Reduction in number of referrals (total - 2WW + 18WW)	19.4%	19.9%	20.4%	32.8%	33.3%	33.8%	43.6%	43.8%	49.0%	49.3%	57.6%	57.8%	-
Reduction in number of referrals (2WW only)	26.5%	28.2%	29.7%	40.7%	42.4%	43.9%	54.2%	55.0%	60.0%	60.7%	69.1%	69.8%	-
Increase in number of referrals (18WW only)	102.7%	122.2%	139.5%	102.7%	122.2%	139.5%	139.5%	147.9%	139.5%	147.9%	139.5%	147.9%	-
Increase in number of repeat FITs	70.2%	79.9%	88.6%	85.4%	95.1%	103.8%	114.9%	119.1%	121.1%	125.3%	130.8%	135.0%	-
Increase in number of watch and waits	37.7%	37.7%	37.7%	68.1%	68.1%	68.1%	90.3%	90.3%	102.7%	102.7%	122.2%	122.2%	-
Reduction in number of COLs	19.4%	19.9%	20.4%	32.7%	33.2%	33.7%	43.4%	43.7%	48.9%	49.1%	57.4%	57.6%	-

Mean time to diagnosis - CRC	2.385	2.419	2.486	2.499	2.529	2.589	2.745	2.800	2.890	2.942	3.294	3.342	1.384
Mean time to diagnosis - AAs	4.678	4.743	4.827	5.268	5.326	5.400	6.015	6.062	6.485	6.530	7.278	7.319	1.956
Mean time to diagnosis - IBD	3.118	3.171	3.249	3.466	3.513	3.583	3.945	3.993	4.224	4.270	4.866	4.908	2.044

▪Southwest quadrant ICER. - AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 61: Tabulated results for QuikRead go

	Int 1: FIT 1 threshold			**Int 2: FIT 2 thresholds**			**Int 3: DG30& NG12**
t (µg/g)	**10**	**100**	**150**	**10/100**	**10/150**	**100/150**	**10**
LYs	14.166	14.163	14.160	14.166	14.166	14.162	14.168
QALYs	10.893	10.890	10.889	10.893	10.892	10.890	10.895
Costs (£)	2,913	2,710	2,692	2,906	2,905	2,710	£3,142
ICER (pairwise, vs Intervention 3)▫ (£)	109,990	97,283	78,165	105,501	101,956	96,085	-
NMB λ=20,000 (vs Int 3) (£)	187	343	335	191	190	342	-
NMB λ=30,000 (vs Int 3) (£)	167	298	277	169	167	297	-
Number of 2WW referrals (total)	0.338	0.126	0.110	0.305	0.302	0.123	0.644
Number of 18WW referrals (total)	0.070	0.092	0.094	0.092	0.094	0.094	0.037
Number of Repeat FITs (total)	0.070	0.092	0.094	0.081	0.082	0.093	0.037
Number of Watch and Wait (total) (total)	0.523	0.690	0.703	0.523	0.523	0.690	0.281
Number of COLs (total)	0.374	0.200	0.187	0.364	0.363	0.200	0.623
Reduction in number of referrals (total - 2WW + 18WW)	40.2%	68.0%	70.1%	41.8%	41.9%	68.2%	-
Reduction in number of referrals (2WW only)	47.5%	80.5%	82.9%	52.7%	53.1%	80.8%	-
Increase in number of referrals (18WW only) ▫▫	85.9%	145.5%	149.9%	-145.5%	-149.9%	-149.9%	-
Reduction in number of COLs	40.0%	67.9%	70.0%	41.7%	41.8%	68.0%	-
Mean time to diagnosis - CRC	2.557	4.438	5.770	2.677	2.762	4.502	1.384
Mean time to diagnosis - AAs	6.418	9.104	9.484	6.586	6.610	9.121	1.956
Mean time to diagnosis - IBD	4.167	6.593	7.027	4.324	4.352	6.615	2.044

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

5.3.8. Conclusions from the cost-effectiveness analyses undertaken in the EAG's base case

The results generated by the EAG indicate that in the vast majority of analyses, the use of FIT has a positive NMB compared with current care. This is produced not by an increase in patient health as there is a very small decrease in estimated QALYs per person (less than 0.005 QALYs per person) but is instead due to the moderate cost savings per person (in the region of £300). This conclusion holds across a wide range of thresholds with the EAG noting that the complex real-world process has been simplified in the model and that uncertainty in parameter inputs results in large uncertainty when directly comparing between thresholds for the same test or comparing directly between tests. The EAG has undertaken sensitivity analysis to explore the robustness of the broad conclusions when using alternative assumptions and data inputs.

The EAG notes that where the use of FIT results in a positive NMB compared with current care, it additionally reduces demand for colonoscopies.

5.3.9. Deterministic scenario analyses

The EAG has run eleven scenario analyses. For illustrative purposes, the sensitivity analyses have all been conducted on the comparison between HM JACKarc using one threshold of 10µg/g (Intervention 1), in comparison to current recommendations (Intervention 3), using the lower intensity option for safety netting. The summary of results is presented in Table 62, whilst full tables are presented in Table 63 to Table 73.

Table 62: Deterministic sensitivity analyses results for HM JACKarc using one threshold (10 µg/g)

	Intervention 1 (FIT using threshold of 10) versus Intervention 3 (DG30/NG12)			
Scenario	**Inc. QALYs**	**Inc. costs**	**ICER▪**	**iNMB (20k)**
Base case (deterministic)	-0.0023	-£327	£143,701	£281
Scenario 1: shorter time to diagnosis (best-case)	-0.0012	-£326	£276,014	£303
Scenario 2: longer time to diagnosis (worst-case)	-0.0041	-£327	£78,942	£245
Scenario 3: QALY loss due to receiving a colonoscopy	-0.0015	-£327	£213,083	£296
Scenario 4: QALY loss for each month of diagnostic delay	-0.0013	-£327	£249,120	£300
Scenario 5: DUAL FIT	-0.0015	-£231	£152,857	£201
Scenario 6: removing IBD and AAs from the model	-0.0012	-£365	£305,949	£341
Scenario 7: Using alternative source for FIT return rate from Moss *et al*. (2017)[125]	-0.0043	-£362	£84,114	£276
Scenario 8: Use of accuracy data for DG30 low-risk group (Intervention 3) from EAG's clinical review analysis for this group	-0.0023	-£289	£125,304	£243
Scenario 9: Increased resource use of GP appointments for patients with NSBP following watch and wait or Repeat FIT	-0.0023	-£314	£138,266	£269
Scenario 10: Alternative method to estimate unit costs for FIT in Intervention 3 (weighted mean)	-0.0023	-£327	£143,919	£282
Scenario 11: FIT has perfect accuracy (sensitivity and specificity =1.0) and return rate =1.0	0.0007	-£441	Dominates	£454

▪*Southwest quadrant ICER*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; QALY - quality-adjusted life year

Scenario 1: shorter time-to-diagnosis (best-case)

Table 63: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.168	14.168	14.168	14.168	14.168	14.167	14.166	14.166	14.165	14.169
QALYs	10.895	10.894	10.894	10.894	10.894	10.894	10.893	10.893	10.893	10.895
Costs (£)	2,954	2,889	2,869	2,842	2,816	2,773	2,730	2,723	2,706	3,142
ICER (pairwise, vs Intervention 3)▫ (£)	264,945	297,382	297,438	290,601	276,014	243,307	196,062	187,295	165,131	-
iNMB λ=20,000 (vs Int 3) (£)	174	237	255	279	303	339	370	374	384	-
iNMB λ=30,000 (vs Int 3) (£)	167	228	245	269	291	324	349	352	357	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)▫▫	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	1.519	1.591	1.626	1.692	1.781	2.018	2.481	2.595	2.947	1.044
Mean time to diagnosis - AAs	2.821	3.194	3.342	3.586	3.881	4.374	5.016	5.141	5.460	1.444
Mean time to diagnosis - IBD	1.863	2.087	2.175	2.324	2.503	2.909	3.501	3.622	3.951	1.396

▪*Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits*
AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 2: longer time to diagnosis (worst-case)

Table 64: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.163	14.163	14.162	14.162	14.161	14.160	14.157	14.157	14.155	14.166
QALYs	10.891	10.890	10.890	10.890	10.889	10.888	10.886	10.886	10.885	10.893
Costs (£)	2,951	2,885	2,866	2,839	2,813	2,769	2,724	2,717	2,698	3,140
ICER (pairwise, vs Intervention 3)▫ (£)	74,399	83,076	83,372	82,131	78,942	71,341	59,436	57,130	51,183	-
iNMB λ=20,000 (vs Int 3) (£)	138	193	208	228	245	267	276	275	269	-
iNMB λ=30,000 (vs Int 3) (£)	113	163	175	191	203	215	206	201	183	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)▫▫	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	4.170	4.449	4.572	4.795	5.087	5.842	7.281	7.630	8.709	2.483
Mean time to diagnosis - AAs	8.151	9.340	9.804	10.565	11.480	13.002	14.971	15.351	16.325	3.702
Mean time to diagnosis - IBD	5.232	5.970	6.253	6.728	7.293	8.556	10.379	10.748	11.753	3.585

▪*Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 3: QALY loss due to receiving a colonoscopy

Table 65: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.892	10.892	10.892	10.892	10.891	10.890	10.890	10.890	10.893
Costs (£)	2953	2888	2868	2841	2815	2772	2728	2721	2703	3,142
ICER (pairwise, vs Intervention 3)▫ (£)	205,902	237,209	236,616	228,623	213,083	180,950	138,858	131,511	113,523	-
NMB λ=20,000 (vs Int 3) (£)	170	233	250	274	296	329	354	357	361	-
NMB λ=30,000 (vs Int 3) (£)	161	222	239	261	281	309	324	325	323	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)▫▫	-72.4%	-92.5%	-98.3%	-106.5%	-114.4%	-127.1%	-139.6%	-141.6%	-146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits
AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 4: QALY loss for each month of diagnostic delay

Table 66: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.856	10.856	10.856	10.856	10.855	10.853	10.848	10.847	10.844	10.893
Costs (£)	2953	2888	2868	2841	2815	2772	2728	2721	2703	3,142
ICER (pairwise, vs Intervention 3)▪ (£)	1,258,362	-1,233,095	-4,870,717	714,990	249,120	102,990	51,135	45,958	35,684	-
NMB λ=20,000 (vs Int 3) (£)	186	258	274	292	300	298	252	238	193	-
NMB λ=30,000 (vs Int 3) (£)	184	260	275	288	287	262	171	146	70	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)▪▪	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

▪Southwest quadrant ICER; ▪▪ Also the value for increased repeat FITs and increased number of watch and waits

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 5: DUAL FIT

Table 67: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold	**Int 3 : DG30& NG12**
t (µg/g)	10	10
LYs	14.167	14.168
QALYs	10.893	10.895
Costs (£)	2,910	3,142
ICER (pairwise, vs Intervention 3)▫ (£)	152,857	-
NMB λ=20,000 (vs Int 3) (£)	201	-
NMB λ=30,000 (vs Int 3) (£)	186	-
Number of 2WW referrals (total)	0.336	0.644
Number of 18WW referrals (total)	0.070	0.037
Number of Repeat FITs (total)	0.070	0.037
Number of Watch and Wait (total) (total)	0.524	0.281
Number of COLs (total)	0.373	0.623
Reduction in number of referrals (total - 2WW + 18WW)	40.4%	-
Reduction in number of referrals (2WW only)	47.7%	-
Increase in number of referrals (18WW only)▫▫	86.3%	-
Reduction in number of COLs	40.2%	-
Mean time to diagnosis - CRC	2.204	1.384
Mean time to diagnosis - AAs	5.512	1.956
Mean time to diagnosis - IBD	2.397	2.044

▪*Southwest quadrant ICER;* ▫▫ *Also the value for increased repeat FITs and increased number of watch and waits*
AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 6: removing IBD and AA from the model

Table 68: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.193	14.193	14.192	14.192	14.192	14.192	14.191	14.190	14.189	14.195
QALYs	10.944	10.944	10.944	10.944	10.944	10.944	10.943	10.943	10.942	10.945
Costs (£)	930	859	839	810	782	754	737	693	686	1,147
ICER (pairwise, vs Intervention 3)▫ (£)	282,336	324,381	326,672	321,193	305,949	281,029	260,920	196,903	185,524	-
NMB λ=20,000 (vs Int 3) (£)	201	270	289	316	341	365	378	408	411	-
NMB λ=30,000 (vs Int 3) (£)	194	261	280	306	329	351	363	385	387	-
Number of 2WW referrals (total)	0.365	0.293	0.272	0.243	0.216	0.189	0.172	0.130	0.124	0.639
Number of 18WW referrals (total)	0.067	0.074	0.077	0.080	0.083	0.085	0.087	0.092	0.092	0.038
Number of Repeat FITs (total)	0.067	0.074	0.077	0.080	0.083	0.085	0.087	0.092	0.092	0.038
Number of Watch and Wait (total) (total)	0.501	0.558	0.575	0.598	0.619	0.641	0.654	0.686	0.692	0.285
Number of COLs (total)	0.394	0.335	0.318	0.295	0.272	0.250	0.237	0.203	0.198	0.617
Reduction in number of referrals (total - 2WW + 18WW)	36.2%	45.8%	48.5%	52.4%	56.0%	59.5%	61.7%	67.2%	68.1%	-
Reduction in number of referrals (2WW only)	42.8%	54.2%	57.5%	62.0%	66.3%	70.5%	73.1%	79.6%	80.6%	-
Increase in number of referrals (18WW only)▫▫	75.9%	96.1%	101.8%	109.9%	117.4%	124.9%	129.6%	141.0%	142.8%	-
Reduction in number of COLs	36.1%	45.7%	48.4%	52.3%	55.9%	59.4%	61.6%	67.1%	68.0%	-
Mean time to diagnosis - CRC	2.296	2.447	2.514	2.638	2.801	3.030	3.224	4.033	4.230	1.384
Mean time to diagnosis - AAs	-	-	-	-	-	-	-	-	-	-
Mean time to diagnosis - IBD	-	-	-	-	-	-	-	-	-	-

▪*Southwest quadrant ICER;* ▫▫ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 7: Using alternative source for FIT return rate from Moss et al. (2017)[125]

Table 69: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.162	14.162	14.162	14.162	14.161	14.161	14.160	14.159	14.159	14.167
QALYs	10.891	10.891	10.891	10.890	10.890	10.890	10.889	10.889	10.888	10.894
Costs (£)	2856	2807	2793	2773	2753	2721	2689	2684	2671	3,115
ICER (pairwise, vs Intervention 3)▫ (£)	70,574	79,586	81,457	83,359	84,114	83,228	77,868	76,448	72,264	-
NMB λ=20,000 (vs Int 3) (£)	185	230	243	260	276	299	316	318	321	-
NMB λ=30,000 (vs Int 3) (£)	149	192	203	219	233	252	262	262	260	-
Number of 2WW referrals (total)	0.296	0.244	0.229	0.207	0.187	0.154	0.121	0.116	0.103	0.622
Number of 18WW referrals (total)	0.074	0.080	0.081	0.083	0.086	0.089	0.093	0.093	0.094	0.040
Number of Repeat FITs (total)	0.074	0.080	0.081	0.083	0.086	0.089	0.093	0.093	0.094	0.040
Number of Watch and Wait (total) (total)	0.555	0.597	0.609	0.626	0.642	0.668	0.694	0.698	0.708	0.299
Number of COLs (total)	0.340	0.297	0.284	0.267	0.250	0.223	0.196	0.192	0.181	0.605
Reduction in number of referrals (total - 2WW + 18WW)	44.0%	51.1%	53.1%	56.0%	58.8%	63.3%	67.7%	68.4%	70.2%	-
Reduction in number of referrals (2WW only)	52.3%	60.8%	63.2%	66.6%	69.9%	75.3%	80.5%	81.4%	83.4%	-
Increase in number of referrals (18WW only)▫▫	85.9%	99.8%	103.8%	109.4%	114.9%	123.6%	132.3%	133.7%	137.0%	-
Reduction in number of COLs	43.9%	51.0%	53.0%	55.9%	58.7%	63.2%	67.6%	68.3%	70.1%	-
Mean time to diagnosis - CRC	4.683	4.787	4.834	4.922	5.038	5.344	5.932	6.075	6.519	1.760
Mean time to diagnosis - AAs	6.285	6.771	6.961	7.275	7.653	8.284	9.104	9.262	9.668	1.956
Mean time to diagnosis - IBD	5.132	5.424	5.537	5.728	5.957	6.470	7.216	7.367	7.779	2.757

▪*Southwest quadrant ICER;* ▫▫ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 8: Use of accuracy data for DG30 low-risk group (Intervention 3) from EAG's clinical review analysis for this group

Table 70: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									**Int 3 : DG30& NG12**
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2953	2888	2868	2841	2815	2772	2728	2721	2703	3104
ICER (pairwise, vs Intervention 3)▫ (£)	106,103	126,902	129,010	128,965	125,304	114,618	96,077	92,392	82,809	-
NMB λ=20,000 (vs Int 3) (£)	123	183	199	222	243	274	298	300	304	-
NMB λ=30,000 (vs Int 3) (£)	108	165	181	202	220	245	259	259	256	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.604
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.042
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.042
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.312
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.591
Reduction in number of referrals (total - 2WW + 18WW)	30.2%	40.1%	43.0%	47.1%	50.9%	57.2%	63.4%	64.4%	66.8%	-
Reduction in number of referrals (2WW only)	42.8%	54.2%	57.5%	62.0%	66.3%	70.5%	73.1%	79.6%	80.6%	-
Increase in number of referrals (18WW only)▫▫	75.9%	96.1%	101.8%	109.9%	117.4%	124.9%	129.6%	141.0%	142.8%	-
Reduction in number of COLs	30.1%	40.0%	42.8%	46.9%	50.7%	57.0%	63.2%	64.2%	66.6%	-
Mean time to diagnosis - CRC	2.325	2.475	2.542	2.665	2.826	3.248	4.055	4.251	4.858	1.375
Mean time to diagnosis - AAs	4.472	5.143	5.405	5.836	6.355	7.220	8.341	8.557	9.112	1.956
Mean time to diagnosis - IBD	2.964	3.370	3.527	3.791	4.105	4.810	5.831	6.038	6.602	2.044

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 9: Increased resource use of GP appointments for patients with NSBP following watch and wait or Repeat FIT

Table 71: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2972	2909	2890	2864	2839	2797	2754	2747	2730	3153
ICER (pairwise, vs Intervention 3)□ (£)	130,551	146,155	146,605	144,199	138,266	124,261	102,703	98,565	87,932	-
NMB λ=20,000 (vs Int 3) (£)	153	211	227	249	269	299	321	323	327	-
NMB λ=30,000 (vs Int 3) (£)	139	194	209	229	246	270	282	282	279	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)□□	-72.4%	-92.5%	-98.3%	-106.5%	-114.4%	-127.1%	-139.6%	-141.6%	-146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

▪Southwest quadrant ICER; □□ Also the value for increased repeat FITs and increased number of watch and waits

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 10: Alternative method to estimate unit costs for FIT in Intervention 3 (weighted mean)

Table 72: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2953	2888	2868	2841	2815	2772	2728	2721	2703	3142
ICER (pairwise, vs Intervention 3)▫ (£)	136,594	152,472	152,841	150,207	143,919	129,194	106,653	102,334	91,243	-
NMB λ=20,000 (vs Int 3) (£)	161	221	238	261	282	313	336	339	343	-
NMB λ=30,000 (vs Int 3) (£)	148	205	220	241	259	284	298	298	295	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)▫▫	-72.4%	-92.5%	-98.3%	-106.5%	-114.4%	-127.1%	-139.6%	-141.6%	-146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 11: FIT with perfect accuracy and return rate=1.0

Table 73: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold	Int 3 : DG30& NG12
t (µg/g)	-	10
LYs	14.169	14.168
QALYs	10.896	10.895
Costs (£)	2711	3,142
ICER (pairwise, vs Intervention 3)▫ (£)	Dominates	-
NMB λ=20,000 (vs Int 3) (£)	454	-
NMB λ=30,000 (vs Int 3) (£)	461	-
Number of 2WW referrals (total)	0.124	0.644
Number of 18WW referrals (total)	0.092	0.037
Number of Repeat FITs (total)	0.092	0.037
Number of Watch and Wait (total) (total)	0.691	0.281
Number of COLs (total)	0.201	0.623
Reduction in number of referrals (total - 2WW + 18WW)	68.6%	-
Reduction in number of referrals (2WW only)	81.0%	-
Increase in number of referrals (18WW only)▫▫	152.1%	-
Reduction in number of COLs	68.1%	-
Mean time to diagnosis - CRC	0.770	1.384
Mean time to diagnosis - AAs	1.668	1.956
Mean time to diagnosis - IBD	0.355	2.044

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits
AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

6 DISCUSSION AND CONCLUSIONS

6.1 Statement of principal findings

6.1.1 Clinical effectiveness – principal findings

The systematic review found no end-to-end RCT studies for any of the tests. The volume of diagnostic test accuracy data differed across tests. Seventeen studies reported across 21 publications[17, 29, 48-51, 53, 55, 59, 63, 64, 66, 69, 70, 74-78, 84, 85] reported data for HM-JACKarc; seventeen studies reported across eighteen publications[32, 43, 45-48, 52, 54, 57, 58, 60, 65, 67, 71, 73, 79-81] reported data for OC-Sensor; three studies[46, 62, 68] reported data for FOB-Gold; one study[61] reported data on QuikRead go; one study[46] reported data for NS-Prime; and one study[82] reported data for both IDK Hb and for IDK Hb/Hp complex. No diagnostic test accuracy data was found for the combined use of IDK Hb + Hb/Hp or for IDK TurbiFIT tests.

Population types: Studies were categorised according to the recruitment criteria as either: population type 1 (studies closest to being a representative spectrum of all patients presenting to primary care with symptoms of CRC who meet NG12 or DG30 criteria); population type 2 (studies closest to being a representative spectrum of NG12 high/medium risk patients); population type 3 (studies closest to being a representative spectrum of DG30 low risk patients); or population type 4 (unclear/likely unrepresentative spectrum). This latter category contained a mixture of studies with different reasons for being "unrepresentative", the most common being that it recruited patients who had been referred to the 2WW secondary care referral pathway. This may be a mixture of NG12 high/medium risk patients and DG30 low risk who were referred from primary care to secondary care on the basis of a positive FIT test in primary care, alongside other patients referred for a variety of reasons. "Enrichment" with patients who were referred on the basis of a positive FIT was thought by the EAG to be a source of heterogeneity between studies that may affect estimates of diagnostic test accuracy (see Section 4.2.1.2).

Main analysis: Since the NICE scope indicated that tests and test-analyser combinations should be considered separately, and since it could not be assumed that all tests were equivalent to each other, the main analysis synthesised data on each test separately. There were only a small number of head-to-head comparative studies and so comparative test accuracy was not formally quantified. Data for Dual FIT were considered separately to single FIT. For each test separately, the main analysis included all population types 1 to 4 together, assuming that the symptoms a patient presents with (i.e., population type) do not affect sensitivity and specificity. Sensitivity analyses were then conducted excluding type 4 studies, which may be enriched with patients who were referred on the basis of a positive FIT or which are otherwise unrepresentative or

unclear. Additional sensitivity analyses were conducted for each of the population types 1, 2 and 3 separately, to see if patient symptoms/population type affected test accuracy. The analyses were only possible for HM-JACKarc and OC-Sensor, due to the small number of studies for the other tests.

The meta-analysis included data at all reported thresholds and provides summary estimates at all possible thresholds. Considering a threshold of 10µg/g, the results were as follows, for sensitivity and specificity respectively: HM-JACKarc (n=16 studies), 89.5% (95% CrI: 84.6,93.4) and 82.8% (95% CrI: 75.2,89.6); OC-Sensor (n=11 studies), 89.8% (95% CrI: 85.9,93.3) and 77.6% (95% CrI: 64.3,88.6); FOB gold (n=3 studies), 91.2% (95% CrI: 68.2,99.8) and 80.3% (95% CrI: 64.9,91.1). No synthesis was conducted for QuikRead go, NS-Prime and IDK tests, since there was only one study for each. For these studies, the estimates of sensitivity and specificity at 10µg/g respectively were: QuikRead go, 92.90% (95% CI: 68.5, 98.7) and 70.10% (95% CI: 66.1, 73.8); and NS-Prime, 71.40% (95% CI: 35.9, 91.8) and 83.60% (95% CI: 78.2, 87.9). The study of IDK Hb and IDK Hb/Hp only reported data at 2µg/g, and the sensitivity and specificity were calculated by IDK to be 87% (95% CI: 84.4, 89.6) and 88.1% (95% CI: 85.6, 90.6); IDK Hb/Hp, 82.6% (95% CI: 79.6, 85.6) and 80.8% (95% CI: 77.7, 83.9). As is usual for diagnostic test accuracy, sensitivity was higher at lower thresholds, and specificity higher at higher thresholds.

The sensitivity analyses showed that the exclusion of type 4 studies did not have significant impact on the pooled estimates, with differences in the point estimates not consistent across the tests, and small in magnitude compared to the uncertainty (as quantified by the credible intervals and prediction intervals). From these analyses, the EAG concludes that it is not necessary to exclude population type 4 studies from the analyses. In the analyses by population types 1, 2 and 3 separately, for HM-JACKarc, the summary sensitivity and specificity for population 3 were higher than for the other population type subgroups, however this analysis was based on only two studies that contributed data at two thresholds (2 and 10µg/g) and was not statistically significantly different based on the overlap of the 95% credible intervals across the subgroups. For the analyses of subgroups by population type for OC-Sensor, the summary estimates were similar and not statistically significant based on overlap of the 95% credible intervals.

Additional analyses 1 and 2: The main analysis was supplemented with two additional analyses. In both additional analyses, all tests were synthesised together to allow the investigation of the impact of study population type and reference standards on a larger sample of studies and because these factors were thought unlikely to be affected by the test type. Additional analysis

1 conducted the same subgroup analyses by population type as described in the previous paragraph, whilst additional analysis 2 restricted to studies where >90% of the patients received a colonoscopy or CTC as the reference standard, to investigate the effect of the reference standard on estimates of diagnostic test accuracy. In additional analysis 2, studies were also subgrouped by test.

In additional analysis 1, in the analyses of subgroups by population type, the summary estimates were similar to the main analysis and not statistically significantly different based on overlap of the 95% CrI.

In additional analysis 2, the summary estimates were similar, irrespective of the reference standard grouping (all studies vs at least 90% of the participants receiving colonoscopy) in all analyses.

Risk of bias assessment: There were risk of bias and/or applicability concerns with all the studies included in the review. Studies mostly fell into two types: a) those that recruited patients referred to secondary care and who had a colonoscopy/CTC/other imaging as the reference standard, as this was part of their routine diagnostic work-up (these are usually population type 2 or 4 studies); and b) those that recruited patients in primary care and for whom the reference standard was either colonoscopy/CTC/other imaging where this was received as part of their diagnostic work up, or was records follow-up where a secondary care referral was not made (these are usually population type 1 or 3 studies). Studies of type a) generally scored high risk of bias for patient selection, since some primary care patients were not recruited (see Section 4.2.1 for a discussion of the different populations), whilst studies of type b) generally scored high risk of bias for the reference standard (see Section 4.1.3 for discussion of the different reference standards), though there were occasional exceptions. Both these factors have been investigated in the statistical synthesis as they could theoretically affect estimates of sensitivity and specificity. Various other sources of bias were noted, for example, the interval between reference standard and index test was poorly reported overall, but due to the "real world" nature of many of the studies is likely to have been within weeks or months rather than years of the index test and is not thought to be a concerning source of bias.

Dual FIT: Four studies reported data using a Dual FIT strategy, two using HM-JACKarc, and one each using OC-Sensor and QuikRead go. In studies that reported estimates for both, the sensitivity was higher and specificity was lower when using Dual FIT (test positive if either FIT positive) than that achieved when using only the first FIT test result to interpret the test.

Comparative diagnostic test accuracy studies: three studies compared two or more tests to each other in the same sample of patients. All three concluded there were some differences between tests, but none were able to conclude whether (and what) different FIT cut-off values would be required for each test. In accordance with this uncertainty about test performance characteristics, the EAG's base case analysis uses data for each test separately.

Patient characteristics subgroup analyses: eleven studies reported diagnostic test accuracy for anaemic patients, three reported data according to age groups, three according to sex, and three for people taking medications which may affect FIT results. No studies were identified according to ethnicity or for people with blood disorders that may affect FIT results. Across these subgroup analyses, evidence was generally limited and sometimes inconsistent. It was not possible to conclude what or whether different FIT thresholds are required according to the patient characteristics specified in the NICE scope.

AA and IBD: Eight studies reported data for the test accuracy of FIT tests for AA and IBD. Uncertainty was high in these analyses, with a large amount of heterogeneity between studies.

Test failures, uptake and repeat tests: Ten studies reported test failure rates and these were largely between 2 and 5%. Only two studies reported test uptake in primary care and only one reported this where return of FIT was part of the diagnostic pathway. In this instance, the non-return rate was 9.4%. For Dual FIT, non-return rates appeared generally higher; all Dual FIT studies were in secondary care.

"Time to" outcomes: data on the time to different points within the diagnostic pathway for patients receiving single or Dual FIT were reported in six diagnostic test accuracy studies but this was largely non-comparative data and difficult to interpret. One further study also reported other outcomes relating to referral rates and emergency presentations and reported reductions in referral rates since introduction of FIT.

Patient perspectives: Two studies reported patient perspectives. The authors conclusions were that most patients found FIT acceptable, but strategies are needed to engage patients with more negative views of FIT, and shared decision making of patient and clinician should be considered for patients dissatisfied with relying on FIT results to decide on need for further investigation. Generalisability of these findings may have been affected by the fact that all patients included had been referred to secondary care.

Sociodemographic factors: One study reported on the impact of sociodemographic factors on FIT return rates and found higher return rates for females compared to males, older patients 65+ years compared to those <65 years, white patients compared to Asian, black and mixed/other ethnic groups, and the least socioeconomically deprived quintile compared to all other quintiles. Suggested strategies for addressing demographic differences in FIT return rate, which may reflect strategies for engagement with services as a whole, included following up after FIT non-return, using multiple languages, shared decision making and patient counselling to address concerns.

6.1.2 Cost-effectiveness – principal findings

The EAG developed a *de novo* health economic model to assess the cost-effectiveness of FIT for people with suspected symptoms of CRC. The model compares three sets of interventions that include the use of quantitative FIT in a primary care setting, exploring a range of different thresholds to determine whether a person would be referred to the 2WW pathway or follow alternative further management pathways (a safety netting pathway or an intermediate group pathway). These latter pathways could result in people being referred to: the 2WW pathway due to ongoing clinical concerns; the 18WW pathway, the watch and wait pathway, or being offered a repeat FIT. The health economic analysis was undertaken from the perspective of the NHS and PSS and was consistent with previous models retrieved by the EAG's review of economic studies, including the one developed to inform NICE DG30.[11] The EAG model adopts a hybrid decision tree – state transition structure. The model parameters were informed by a number of sources including the EAG's clinical review and synthesis, NCRAS and ONS datasets, previous NICE TAs (TA856 and TA342), expert clinical opinion, the Whyte et al. (2023) model, and standard costing sources.

The EAG's base case model suggests that for all FIT brands there are strategies with a positive iNMB compared with current care regardless of the cost-effectiveness threshold used, or whether one or two thresholds were used. This was due to cost savings associated with reduced colonoscopies although this was at the expense of a slight reduction in patient health caused by patients who previously would have had a colonoscopy receiving a false negative FIT result. These conclusions produced by the EAG's base case analysis were robust to the sensitivity analyses undertaken.

The exact brand and threshold(s) which generate the greatest iNMB (at a selected threshold) could not be robustly determined due to the similarity of iNMB values, parameter uncertainty and the possibility of omissions from the model structure.

6.2 Strengths and limitations of the assessment

6.2.1 Strengths and limitations in the clinical evidence base

6.2.1.1 Strengths and limitations of the evidence base

Whilst the evidence base was large, it was also complicated and incomplete with respect to the scope issued by NICE. Key evidence gaps were for the IDK tests. No diagnostic test accuracy data was found for IDK Hb + Hb/Hp, since the EAG was of the opinion that an assumption of independence between the two tests that make up this test could not be made. No diagnostic test accuracy data was identified for the IDK TurbiFIT tests, and the EAG was of the opinion that the analysis comparing IDK TurbiFIT to IDK Hb was not sufficient for the EAG to make a formal recommendation on equivalence. Additionally, data for IDK Hb and Hb/Hp complex came from the one small study with limited details about patient recruitment, and no data was provided to show that the test used in that study was equivalent to the current commercial test. Similarly, evidence was limited for NS-Prime (1 study, n recruited =233, n of CRC events = 7) and QuikRead go (n recruited =553, n of CRC events = 14). The NS-Prime study was conducted in a subgroup of patients from the NICE FIT study who returned all 4 tests, which may have introduced additional generalisability concerns if non-return of FIT meant the patient spectrum was altered in such a way that may affect the estimates of sensitivity or specificity, e.g., excluded older age. No diagnostic test accuracy data was found relating to OC-Sensor Ceres, and the correlation data provided by the company comparing OC-Sensor Ceres to OC-Sensor iO and OC-Sensor PLEDIA was not sufficient for the EAG to make a formal recommendation on equivalence. Reporting of the inclusion or exclusion of patients with "bypass" symptoms (rectal or anal mass or anal ulceration) was often missing from the studies, and some studies excluded rectal bleeding, factors which may affect the patient spectrum in comparison to the scope. Data on test failures, test uptake and repeat tests were largely only available for HM-JACKarc and OC-Sensor. There was also no diagnostic test data according to ethnicity or for people with blood disorders that may affect FIT, and the available data on other patient characteristics was not conclusive. Data on patient outcomes such as HRQoL and anxiety were not available in the studies that reported diagnostic test accuracy and it was beyond the scope of this assessment and time available to review this data in other study designs.

6.2.1.2 Strengths and limitations of the systematic review

The systematic review was conducted to high standards and used two reviewers to validate data extractions and risk of bias assessments. However, there were limitations due to the limited time available to complete the work, combined with a large and complicated evidence base. Amongst the limitations were the use of one reviewer to conduct most of the study selection process, which may have resulted in studies being missed, though potential errors in misunderstanding the inclusion criteria were mitigated by concordance between two reviewers

being established on the first 200 records. Clinical advisors and Specialist Committee Members were also consulted for potentially missed studies. Studies may also have been missed if these were excluded by the original DG30 review, though this is thought unlikely since that review had wide inclusion criteria. An additional element to the search for this assessment was also added to mitigate against missed studies, by including search terms for each of the tests, without date limits. In addition, some quality assessment work could not be completed in the time available, though all studies that contributed to the three analyses this affected (comparative diagnostic test accuracy studies; Dual FIT studies; AA and IBD studies) were assessed in the context of the other analyses they contributed to, except for one (Tsapournas 2020).[83]

Due to the emphasis in this project to identify the optimal way to use FIT to reduce the number of people without significant bowel pathology who are referred to the suspected CRC pathway, taking into consideration the threshold used to define a positive test, the synthesis used an advanced statistical model that accommodates estimates of sensitivity and specificity at multiple thresholds from each study. This has several advantages over the more commonly used approach of performing separate bivariate meta-analyses at selected thresholds, including making use of all available data, increasing precision, ensuring consistency of pooled results, and producing summary estimates at all thresholds of interest to be considered in the cost-effectiveness model. However, some of the analyses are subject to considerable uncertainty due to the small number of studies and should be interpreted with caution. There were several potential sources of between study heterogeneity. Although these were explored using subgroup analyses it was not possible to make conclusive recommendations on any of these factors. Although it would be possible to extend the presented synthesis to include covariates that may explain the heterogeneity between studies (such as population type, reference standard, population characteristics) this was not conducted due to time constraints, and challenges presented by low numbers of studies in certain subgroups.

It was challenging to assign studies to population-type categories and whilst authors were contacted to clarify inclusion criteria, this did not always resolve ambiguities. As such, it is possible that some studies have been wrongly categorised by the EAG reviewers. This may also have affected the sensitivity analyses done to test the effect of population type, since if studies have been miscategorised, this may have altered the effect of removing them and obscured real differences.

It is thought likely by clinical advisors to the EAG that FIT tests will be used in a wider spectrum of patients (including those with less serious symptoms) in primary care than only those with NG12 high/medium-risk or DG30 low-risk symptoms. It is unclear if test accuracy

would be similar in a wider spectrum of patients with less serious symptoms. It was not the focus of this assessment to consider this issue, since the scope was limited to NG12 and DG30 patients. Three studies[29, 44, 66] were highlighted by clinical advisors to the EAG as having potentially recruited a wider population than just NG12 or DG30 patients. In our analysis, two of these studies[29, 44, 66] contributed to the type 4 subgroup, since they were not exclusively DG30 low-risk patients, whilst the other[44] was categorised as a type 1 study. It is possible that some of the other studies, in particular those in Scotland, also recruited wider populations.

Due to time constraints, it was not possible to consider the impact of distribution and sample return methods on return rates. Methods encountered in the literature included distribution by GPs in a face-to-face appointment, distribution by post, or distribution in secondary care. Other methods may be used across the country. It was also not possible to consider the causes of test failures. Causes reported in the literature include labelling errors, incorrect containers, no date of collection or sample too old, volume errors and laboratory accidents. These may be amenable to improvement through training of GPs (e.g. in how to describe the test to patients), patient information leaflets (e.g. to avoid overfilling and labelling/date errors) and laboratory personnel (e.g. in how to avoid accidents), or other interventions to avoid test failures.

6.2.1.3 Comparison to other analyses

This analysis has some differences in estimates compared to the BSG/ACPGBI review, and the DG30 review. At a threshold of 10µg/g, the BSG/ACPGBI review found a pooled sensitivity and specificity respectively for HM-JACKarc of 95.2% (95% CI: 86.5, 99.0) and 78.2% (95% CI: 69.2, 85.2) compared to 89.5% (95% CrI: 84.6,93.4) and 82.8% (95% CrI: 75.2,89.6) in the EAG's analysis, and 100% (95% CI 71.5–100%) and 76.6% (95% CI 72.6–80.3%) in the DG30 analysis. For OC-Sensor the ACPGBI/BSG analysis pooled estimates were 90.2% (95% CI: 86.2, 93.1) and 74.5% (95% CI: 68.1, 79.9), compared to 89.8% (95% CrI: 85.9,93.3) and 77.6% (95% CrI: 64.3,88.6) in the EAG's analysis, and 92.1% (9CI: CI 86.9–95.3%), and specificity was 85.8% (95% CI 78.3–91.0%) in the DG30 analysis. For FOB gold the ACPGBI/BSG analysis pooled estimates were 95.2% (95% CI: 86.5, 99.0) and 71.3% (95% CI: 68.0, 74.3), compared to 91.2% (95% CrI: 68.2,99.8) and 80.3% (95% CrI: 64.9,91.1) in the EAG's analysis (there was no FOB Gold data in the DG30 review).

The differences in estimates are generally small and may be due to the relatively large number of additional studies and patients included in the review for this assessment compared to the DG30 and ACPGBI/BSG review, even though less than a year had elapsed since the ACPGBI/BSG review searches were conducted. A bivariate meta-analysis including 14 HM-JACKarc studies that report diagnostic test accuracy at a threshold of 10 was conducted by the

EAG and found a pooled sensitivity and specificity of 89.2 (95% CrI: 85.7, 92.0) and 79.4 (95% CrI: 75.0, 83.3) respectively. This suggests that the larger sensitivity reported in previous reviews may be largely explained by the difference in studies contributing to the analysis, rather than the different statistical methods used (stratified bivariate model vs multiple thresholds model). The difference in the specificity is within the confidence intervals reported across the analyses and may be due to the different studies that have entered the analysis, and/or the methods of the multiple threshold model.

6.2.2 Strengths and limitations relating to the health economic analysis

The EAG model has a number of strengths; in particular: (i) the model structure builds upon other published models that evaluates FIT in people with symptoms of CRC; (ii) the model includes colonoscopy capacity to impact on the waiting times for patients following the 2WW and 188WW pathways; (iii) the inclusion of AAs and IBD within the mathematical model; and (iv) the uncertainty in the model inputs and assumptions has been explored in sensitivity analyses.

However, the model is also subject to several limitations and uncertainties relating to the cost-effectiveness analysis, which include: (i) the uncertainty in data inputs, particularly diagnostic accuracy data and those reliant on expert opinion; (ii) the structure of the model may have omitted aspects of the complex real-world problem; and (iii) the relative similarity in iNMB values for FIT strategies meant that no robust estimate of the FIT brand or the threshold(s) which generated the greatest iNMB could be made.

6.3 Uncertainties

It was beyond the scope of this assessment to conduct cost-effectiveness analyses for the patient characteristics subgroups defined in the scope, and clinical data limitations would have prevented such analyses had they been planned. The project deadlines and a lack of evidence also prevented a more in-depth analyses of the inputs informed by clinical opinion.

Evidence for the accuracy of IDK TurbiFit was lacking; therefore, no analyses could be made for this test. Diagnostic test accuracy data for four other tests (QuikRead go, NS-Prime, IDK Hb and IDK Hb/Hp) were reliant on a single fairly small study each.

The EAG also notes that the standard care of the model may not reflect current use of the test in some locations in England, as it is known that there is some heterogeneity between diagnostic and clinical management of patients with suspected CRC. However, current care was intended

to reflect NICE current recommendations as defined in DG30 and NG12, in accordance with the NICE scope.

6.4 Generalisability

The assessment included only studies that were conducted in patients who presented to primary care with symptoms of CRC, except in cases where insufficient evidence necessitated the use of studies that also recruited from secondary care. This affected the Dual FIT QuikRead go analysis and the analysis of medications that might affect FIT results. As noted, the assessment of the evidence base has considered potential sources of heterogeneity in terms of the populations recruited to the included studies, and whilst analyses did not indicate population type affected estimates, the limitations of the analyses (e.g., difficulties categorising studies) mean they were not conclusive.

Data on test failures, test uptake and repeat tests were largely only available for HM-JACKarc and OC-Sensor; the generalisability of these data to other tests has been assumed in the model. For Dual FIT for these outcomes, data were only available from studies conducted in secondary care, which may affect generalisability.

As noted, heterogeneity in clinical practice across England may affect the generalisability of some modelling assumptions, e.g., safety netting. The model was robust to all scenario analyses.

6.5 Implications for service provision

The model makes assumptions about the effects of safety netting, which may not be consistent with the safety netting offered across the country at present. Standardisation of and improvement to safety netting practice may be required. Interventions may be required to increase FIT return, especially in some socioeconomic groups, and to improve the experience of a minority who have negative views about FIT, and dissatisfaction with reliance on FIT for diagnostic purposes. The optimal way to distribute FIT should be considered, e.g. via post or via the GP. Ways to avoid test failures should also be considered, e.g. GP training, patient information leaflets, and laboratory staff training. Implementation of FIT in patients with symptoms defined in NG12 and DG30 may lead to use of FIT in a wider group of patients, and this possibility may need to be monitored and/or mitigated against.

6.6 Suggested research priorities

The comparative diagnostic test accuracy between tests remains uncertain, largely due to the limited evidence base comparing tests to one another; new primary research studies may be

required. It remained unclear whether and what different thresholds are required for patients with characteristics that may affect FIT accuracy; new primary research studies may be required. Whilst the analysis was not able to detect an effect of population type, enrichment with FIT positives or the reference standard used, these are all issues that should be considered in future primary studies and evidence syntheses since the analyses conducted here were not conclusive. Efforts could be made to include the relevant patient spectrum with the best possible reference standard, though it is unlikely that it would be appropriate to give all patients presenting to primary care a colonoscopy due to costs, risk of adverse events and patient preferences with respect to this test. The issues described in this report relating to the impact of the reference standard and population type on test accuracy estimates are likely to remain difficult to disentangle. It is unclear if test accuracy would be similar in a wider spectrum of patients with less serious symptoms, and future primary studies or a careful consideration of the existing evidence base, as noted above in Section 4.4.3, could address this issue.

6.7 The use of patient and public involvement

There was no patient and public involvement in producing this report. This was not considered possible within the timescales of the project. However, the EAG is aware that at the NICE Technology Appraisal Committee that will discuss this topic, there will be patient and public involvement and representation, and this may result in the EAG changing model parameters and generating revised results.

6.8 Equality, Diversity, and Inclusion

As this report is secondary research, no patient participation was involved and the EAG did not need to consider the equality, diversity, and inclusion of participants. The primary research team was part of the ScHARR Technology Assessment Group contracted by the Department of Health, and this team is a group representing a range of protected characteristics, consisting of seniority, ages, ethnicity, and religious beliefs, and including both males and females. The lead author is not the most senior member of the team.

7 REFERENCES

1. Jasperson KW, Tuohy TM, Neklason DW, Burt RW. Hereditary and familial colon cancer. *Gastroenterology* 2010;138:2044-58.
2. Keum N, Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nature reviews Gastroenterology & hepatology* 2019;16:713-32.
3. Bretthauer M, Kalager M, Adami H-O. Do's and don'ts in evaluation of endoscopic screening for gastrointestinal cancers. *Endoscopy* 2015:75-80.
4. Cancer Research UK. Bowel cancer statistics. 2022. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer (Accessed
5. IHME GHDx. GBD Results. 2023.
6. Vuik FE, Nieuwenburg SA, Bardou M, Lansdorp-Vogelaar I, Dinis-Ribeiro M, Bento MJ*, et al.* Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. *Gut* 2019;68:1820-6.
7. Oliphant R, Brewster D, Morrison D. The changing association between socioeconomic circumstances and the incidence of colorectal cancer: a population-based study. *British journal of cancer* 2011;104:1791-6.
8. National Cancer Intelligence Network. Cancer by deprivation in England: Incidence, 1996–2010 & Mortality, 1997–2011. *Journal* 2014.
9. Tweed E, Allardice G, McLoone P, Morrison D. Socio-economic inequalities in the incidence of four common cancers: a population-based registry study. *Public health* 2018;154:1-10.
10. NICE. Suspected cancer: recognition and referral. 2015.
11. NICE. Quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care. Diagnostics guidance [DG30]; 2017.
12. Monahan KJ, Davies MM, Abulafi M, Banerjea A, Nicholson BD, Arasaradnam R*, et al.* Faecal immunochemical testing (FIT) in patients with signs or symptoms of suspected colorectal cancer (CRC): a joint guideline from the Association of Coloproctology of Great Britain and Ireland (ACPGBI) and the British Society of Gastroenterology (BSG). *Gut* 2022;71:1939.
13. . Using Faecal Immunochemical Testing (FIT) in the Lower Gastrointestinal (GI) pathway (primary care). London; 2022.
14. . Using Faecal Immunochemical Testing (FIT) in the Lower Gastrointestinal (GI) pathway (secondary care). London; 2022.
15. Mozdiak E, Weldeselassie Y, McFarlane M, Tabuso M, Widlak MM, Dunlop A*, et al.* Systematic review with meta-analysis of over 90 000 patients. Does fast-track review diagnose colorectal cancer earlier? *Alimentary pharmacology & therapeutics* 2019;50:348-72.
16. Booth R, Carten R, D'Souza N, Westwood M, Kleijnen J, Abulafi M. Role of the faecal immunochemical test in patients with risk-stratified suspected colorectal cancer symptoms: A systematic review and meta-analysis to inform the ACPGBI/BSG guidelines. *The Lancet Regional Health-Europe* 2022;23:100518.
17. D'Souza N, Delisle TG, Chen M, Benton S, Abulafi M. Faecal immunochemical test is superior to symptoms in predicting pathology in patients with suspected colorectal cancer symptoms referred on a 2WW pathway: a diagnostic accuracy study. *Gut* 2021c;70:1130-8.
18. D'Souza N, Hicks G, Benton S, Abulafi M. The diagnostic accuracy of the faecal immunochemical test for colorectal cancer in risk-stratified symptomatic patients. *The Annals of The Royal College of Surgeons of England* 2020;102:174-9.
19. Fraser CG, Allison JE, Young GP, Halloran SP, Seaman HE. Improving the reporting of evaluations of faecal immunochemical tests for haemoglobin. *European Journal of Cancer Prevention* 2015;24:24-6.

20. Carroll M, Piggott C, Pearson S, Seaman H, Bruce H, Halloran S. AN EVALUATION OF QUANTITATIVE FAECAL IMMUNOCHEMICAL TESTS FOR HAEMOGLOBIN. GUT, abstract no. 4236, p. A129-A30.
21. NICE. Quantitative faecal immunochemical tests to guide referral in primary care for people with high risk symptoms. Final scope. 2020.
22. NICE. Quantitative faecal immunochemical tests to guide colorectal cancer pathway referral in primary care. Final scope. 2022.
23. Farkas NG, Fraser CG, Maclean W, Jourdan I, Rockall T, Benton SC. Replicate and repeat faecal immunochemical tests in symptomatic patients: A systematic review. *Annals of Clinical Biochemistry* 2023;60:27-36.
24. Westwood M, Corro Ramos I, Lang S, Luyendijk M, Zaim R, Stirk L*, et al.* Faecal immunochemical tests to triage patients with lower abdominal symptoms for suspected colorectal cancer referrals in primary care: A systematic review and costeffectiveness analysis. 2016.
25. Monahan KJ, Davies MM, Abulafi M, Banerjea A, Nicholson BD, Arasaradnam R*, et al.* Faecal immunochemical testing (FIT) in patients with signs or symptoms of suspected colorectal cancer (CRC): a joint guideline from the Association of Coloproctology of Great Britain and Ireland (ACPGBI) and the British Society of Gastroenterology (BSG). *Gut* 2022;12:12.
26. Booth R, Carten R, D'Souza N, Westwood M, Kleijnen J, Abulafi M. Role of the faecal immunochemical test in patients with risk-stratified suspected colorectal cancer symptoms: A systematic review and meta-analysis to inform the ACPGBI/BSG guidelines. *The Lancet Regional Health Europe* 2022;23:100518.
27. Burr NE, Derbyshire E, Taylor J, Whalley S, Subramanian V, Finan PJ*, et al.* Variation in post-colonoscopy colorectal cancer across colonoscopy providers in English National Health Service: population based cohort study. *bmj* 2019;367.
28. Obaro AE, Plumb AA, Fanshawe TR, Torres US, Baldwin-Cleland R, Taylor SA*, et al.* Post-imaging colorectal cancer or interval cancer rates after CT colonography: a systematic review and meta-analysis. *The Lancet Gastroenterology & Hepatology* 2018;3:326-36.
29. Nicholson BD, James T, Paddon M, Justice S, Oke JL, East JE*, et al.* Faecal immunochemical testing for adults with symptoms of colorectal cancer attending English primary care: a retrospective cohort study of 14 487 consecutive test requests. *Alimentary Pharmacology & Therapeutics* 2020;52:1031-41.
30. Centre for Reviews and Dissemination. CRD's guidance for undertaking reviews in health care. 2008.
31. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Cochrane handbook for systematic reviews of diagnostic test accuracy. In: Version; 2010.
32. Crooks C, Banerjea A, Jones J, Chapman C, Oliver S, West J*, et al.* Assessing empirical thresholds for investigation in people referred on a symptomatic colorectal cancer pathway: a cohort study utilising faecal immunochemical and blood tests in England. *medRxiv* 2023; 10.1101/2023.03.29.23287919:2023.03.29.23287919.
33. Bailey J, Morton A, Jones J, Chapman C, Oliver S, Morling J*, et al.* Sociodemographic Variations in the Uptake of Faecal Immunochemical Tests in Primary Care. In: 2023 StN, ed.; 2023a.
34. Georgiou Delisle T, D'Souza N, Davies B, Benton S, Chen M, Ward H*, et al.* Faecal immunochemical test for suspected colorectal cancer symptoms: patient survey of usability and acceptability. *BJGP Open* 2022b;6.
35. Maclean W, Whyte MB, Farkas N, Benton SC, Rockall T, Jourdan I. Patient-reported outcome measures show FIT as an acceptable investigation to rule out colorectal cancer in the two-week wait cohort. *Annals of the Royal College of Surgeons of England* 2022b;31.
36. Whiting P, Rutjes A, Westwood M, Mallett S, Deeks J, Reitsma J*, et al.* QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 2011;155:529-36.

37. Yang B, Mallett S, Takwoingi Y, Davenport CF, Hyde CJ, Whiting PF*, et al.* QUADAS-C: a tool for assessing risk of bias in comparative diagnostic accuracy studies. *Annals of internal medicine* 2021;174:1592-9.
38. Jones HE, Gatsonsis CA, Trikalinos TA, Welton NJ, Ades AE. Quantifying how diagnostic test accuracy depends on threshold in a meta-analysis. *Statistics in Medicine* 2019;38:4789-803.
39. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. In. Vienna, Austria; 2022.
40. Plummer M. rjags: Bayesian Graphical Models using MCMC. R package version 4-10. In; 2019.
41. Brooks S, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998;7:434-55.
42. Spiegelhalter D, Best N, Carlin B, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B Stat Methodol* 2002;64:583-639.
43. Archer T, Aziz I, Kurien M, Knott V, Ball A. Prioritisation of lower gastrointestinal endoscopy during the COVID-19 pandemic: outcomes of a novel triage pathway. *Frontline Gastroenterology* 2022;13:225-30.
44. Bailey JA, Weller J, Chapman CJ, Ford A, Hardy K, Oliver S*, et al.* Faecal immunochemical testing and blood tests for prioritization of urgent colorectal cancer referrals in symptomatic patients: A 2-year evaluation. *BJS Open* 2021a;5(2) (no pagination).
45. Ball AJ, Aziz I, Parker S, Sargur RB, Aldis J, Kurien M. Fecal Immunochemical Testing in Patients With Low-Risk Symptoms of Colorectal Cancer: A Diagnostic Accuracy Study. *Journal of the National Comprehensive Cancer Network* 2022;20:989-96.e1.
46. Benton SC, Piggott C, Zahoor Z, O'Driscoll S, Fraser CG, D'Souza N*, et al.* A comparison of the faecal haemoglobin concentrations and diagnostic accuracy in patients suspected with colorectal cancer and serious bowel disease as reported on four different faecal immunochemical test systems. *Clinical Chemistry & Laboratory Medicine* 2022;60:1278-86.
47. Cama R, Kapoor N, Sawyer P, Patel B, Landy J. Evaluation of 13,466 Fecal Immunochemical Tests in Patients Attending Primary Care for High- and Low-Risk Gastrointestinal Symptoms of Colorectal Cancer. *Digestive Diseases & Sciences* 2022;10:10.
48. Chapman CJ, Banerjea A, Humes DJ, Allen J, Oliver S, Ford A*, et al.* Choice of faecal immunochemical test matters: comparison of OC-Sensor and HM-JACKarc, in the assessment of patients at high risk of colorectal cancer. *Clinical Chemistry & Laboratory Medicine* 2021;59:721-8.
49. D'Souza N, Hicks G, Benton SC, Abulafi M. The diagnostic accuracy of the faecal immunochemical test for colorectal cancer in risk-stratified symptomatic patients. *Annals of the Royal College of Surgeons of England* 2020a;102:174-9.
50. Elbeltagi A, Salama M, Boxall P, Roos J, Lim M. The Yield of Faecal Immunochemical Test in the Detection of Colorectal Cancer within a Fast-track Pathway at York, United Kingdom. *Turkish Journal of Colorectal Disease* 2022;32(3):178-85.
51. Faux JW, Cock K, Bromley R, Feldman M. Colorectal two-week wait service and quantitative FIT: it's not just about colon cancer. *Annals of the Royal College of Surgeons of England* 2022;104:257-60.
52. Georgiou Delisle T, D'Souza N, Tan J, Najdawi A, Chen M, Ward H*, et al.* Introduction of an integrated primary care faecal immunochemical test referral pathway for patients with suspected colorectal cancer symptoms. *Colorectal Disease* 2022a;08:08.
53. Godber IM, Todd LM, Fraser CG, MacDonald LR, Younes HB. Use of a faecal immunochemical test for haemoglobin can aid in the investigation of patients with

lower abdominal symptoms. *Clinical Chemistry & Laboratory Medicine* 2016;54:595-602.
54. Hunt N, Rao C, Logan R, Chandrabalan V, Oakey J, Ainsworth C*, et al.* A cohort study of duplicate faecal immunochemical testing in patients at risk of colorectal cancer from North-West England. *BMJ Open* 2022;12:e059940.
55. Johnstone MS, Burton P, Kourounis G, Winter J, Crighton E, Mansouri D*, et al.* Combining the quantitative faecal immunochemical test and full blood count reliably rules out colorectal cancer in a symptomatic patient referral pathway. *International Journal of Colorectal Disease* 2022a;37:457-66.
56. Johnstone MS, MacLeod C, Digby J, Al-Azzawi Y, Pang G, Watson AJM*, et al.* Prevalence of repeat faecal immunochemical testing in symptomatic patients attending primary care. *Colorectal Disease* 2022b;01:01.
57. Juul JS, Hornung N, Andersen B, Laurberg S, Olesen F, Vedsted P. The value of using the faecal immunochemical test in general practice on patients presenting with non-alarm symptoms of colorectal cancer. *British Journal of Cancer* 2018;119(4):471-9.
58. Laszlo HE, Seward E, Ayling RM, Lake J, Malhi A, Stephens C*, et al.* Faecal immunochemical test for patients with 'high-risk' bowel symptoms: a large prospective cohort study and updated literature review. *British Journal of Cancer* 2022;126:736-43.
59. MacDonald S, MacDonald L, Godwin J, Macdonald A, Thornton M. The diagnostic accuracy of the faecal immunohistochemical test in identifying significant bowel disease in a symptomatic population. *Colorectal Disease* 2022;24:257-63.
60. Maclean W, Limb C, Mackenzie P, Whyte MB, Benton SC, Rockall T*, et al.* Adoption of faecal immunochemical testing for 2-week-wait colorectal patients during the COVID-19 pandemic: an observational cohort study reporting a new service at a regional centre. *Colorectal Disease* 2021a;23(7):1622-9.
61. Maclean W, Mackenzie P, Limb C, Zahoor Z, Whyte MB, Rockall T*, et al.* Diagnostic accuracy of point of care faecal immunochemical testing using a portable high-speed quantitative analyser for diagnosis in 2-week wait patients. *Colorectal Disease* 2021b;23:2376-86.
62. MacLean W, Zahoor Z, O'Driscoll S, Piggott C, Whyte MB, Rockall T*, et al.* Comparison of the QuikRead gopoint-of-care faecal immunochemical test for haemoglobin with the FOB Gold Widelaboratory analyser to diagnose colorectal cancer in symptomatic patients. *Clinical Chemistry and Laboratory Medicine* 2022a;60(1):101-8.
63. Mowat C, Digby J, Strachan JA, McCann R, Hall C, Heather D*, et al.* Impact of introducing a faecal immunochemical test (FIT) for haemoglobin into primary care on the outcome of patients with new bowel symptoms: a prospective cohort study. *BMJ Open Gastroenterology* 2019;6:e000293.
64. Mowat C, Digby J, Strachan JA, McCann RK, Carey FA, Fraser CG*, et al.* Faecal haemoglobin concentration thresholds for reassurance and urgent investigation for colorectal cancer based on a faecal immunochemical test in symptomatic patients in primary care. *Annals of Clinical Biochemistry* 2021;58:211-9.
65. Mowat C, Digby J, Strachan JA, Wilson R, Carey FA, Fraser CG*, et al.* Faecal haemoglobin and faecal calprotectin as indicators of bowel disease in patients presenting to primary care with bowel symptoms. *Gut* 2016;65:1463-9.
66. Nicholson BD, James T, East JE, Grimshaw D, Paddon M, Justice S*, et al.* Experience of adopting faecal immunochemical testing to meet the NICE colorectal cancer referral criteria for low-risk symptomatic primary care patients in Oxfordshire, UK. *Frontline Gastroenterology* 2019;10:347-55.
67. Pin-Vieito N, Garcia Nimo L, Bujanda L, Roman Alonso B, Gutierrez-Stampa MA, Aguilar-Gama V*, et al.* Optimal diagnostic accuracy of quantitative faecal immunochemical test positivity thresholds for colorectal cancer detection in primary

health care: A community-based cohort study. *United European Gastroenterology Journal* 2021;9:256-67.
68. Schwettmann L, Lied A, Eriksen R. Evaluation of the Sentinel-FOB gold faecal immunochemical test for the presence of haemoglobin using the automated Roche Cobas 8000 system. *Practical Laboratory Medicine* 2022;29:e00263.
69. Tang A, Chandler S, Torkington J, Harris DA, Dhruva Rao PK. Adapting the investigation of patients on urgent suspected cancer pathway with lower gastrointestinal symptoms across Wales during COVID-19. *Annals of the Royal College of Surgeons of England* 2022;26:26.
70. Withrow DR, Shine B, Oke J, Tamm A, James T, Morris E*, et al.* Combining faecal immunochemical testing with blood test results for colorectal cancer risk stratification: a consecutive cohort of 16,604 patients presenting to primary care. *BMC Medicine* 2022;20:116.
71. Ayling RM, Lewis SJ, Cotter F. Potential roles of artificial intelligence learning and faecal immunochemical testing for prioritisation of colonoscopy in anaemia. *British Journal of Haematology* 2019;185:311-6.
72. Bailey J, Morton A, Jones J, Chapman C, Oliver S, Morling J*, et al.* "Low FIT" Colorectal cancer: A four-year comparison of the Nottingham "4F" protocol with FIT10 in symptomatic patients. In; 2023b.
73. Bujanda L, Sarasqueta C, Vega P, Salve M, Quintero E, Alvarez-Sánchez V*, et al.* Effect of aspirin on the diagnostic accuracy of the faecal immunochemical test for colorectal advanced neoplasia. *United European Gastroenterol J* 2018;6:123-30.
74. Cunin L, Khan AA, Ibrahim M, Lango A, Klimovskij M, Harshen R. FIT negative cancers: A right-sided problem? Implications for screening and whether iron deficiency anaemia has a role to play. *The Surgeon* 2021;19:27-32.
75. D'Souza N, Delisle TG, Chen M, Benton SC, Abulafi M, the NFITSC. Faecal immunochemical testing in symptomatic patients to prioritize investigation: diagnostic accuracy from NICE FIT Study. *British Journal of Surgery* 2021a;108:804-10.
76. D'Souza N, Monahan K, Benton SC, Wilde L, Abulafi M, Group NFS*, et al.* Finding the needle in the haystack: the diagnostic accuracy of the faecal immunochemical test for colorectal cancer in younger symptomatic patients. *Colorectal Disease* 2021b;23:2539-49.
77. Farrugia A, Widlak M, Evans C, Smith SC, Arasaradnam R. Faecal immunochemical testing (FIT) in symptomatic patients: what are we missing? *Frontline Gastroenterol* 2020;11:28-33.
78. Gerrard AD, Maeda Y, Miller J, Gunn F, Theodoratou E, Noble C*, et al.* Double faecal immunochemical testing in patients with symptoms suspicious of colorectal cancer. *British Journal of Surgery* 2023;110:471-80.
79. Morales Arraez D, Carrillo G, Adrian M, Gimeno Z, Quintero A. Role of faecal immunochemical testing in the diagnostic workup of patients with iron deficiency anaemia. *United Eur Gastroenterol J* 2018;6:A403–A4.
80. Rodriguez-Alonso L, Rodriguez-Moranta F, Arajol C, Gilabert P, Serra K, Martin A*, et al.* Proton pump inhibitors reduce the accuracy of faecal immunochemical test for detecting advanced colorectal neoplasia in symptomatic patients. *PLoS One* 2018;13:e0203359.
81. Rodriguez-Alonso L, Rodriguez-Moranta F, Ruiz-Cerulla A, Arajol C, Serra K, Gilabert P*, et al.* The use of faecal immunochemical testing in the decision-making process for the endoscopic investigation of iron deficiency anaemia. *Clin Chem Lab Med* 2020;58:232-9.
82. Sieg A, Thoms C, Lüthgens K, John MR, Schmidt-Gayk H. Detection of colorectal neoplasms by the highly sensitive hemoglobin-haptoglobin complex in feces. *International Journal of Colorectal Disease* 1999;14:267-71.
83. Tsapournas G, Hellström PM, Cao Y, Olsson LI. Diagnostic accuracy of a quantitative faecal immunochemical test vs. symptoms suspected for colorectal

cancer in patients referred for colonoscopy. *Scandinavian Journal of Gastroenterology* 2020;55:184-92.
84. Turvill J, Mellen S, Jeffery L, Bevan S, Keding A, Turnock D. Diagnostic accuracy of one or two faecal haemoglobin and calprotectin measurements in patients with suspected colorectal cancer. *Scandinavian Journal of Gastroenterology* 2018;53:1526-34.
85. Turvill J, Turnock D, Cottingham D, al. e. The Fast Track FIT study: diagnostic accuracy of faecal immunochemical test for haemoglobin in patients with suspected colorectal cancer. *Br J Gen Pract* 2021;71:E643–E51.
86. Cubiella J, Salve M, Díaz-Ondina M, Vega P, Alves MT, Iglesias F*, et al.* Diagnostic accuracy of the faecal immunochemical test for colorectal cancer in symptomatic patients: comparison with NICE and SIGN referral criteria. *Colorectal Dis* 2014;16:O273-82.
87. Eiken Chemical Co. L. Analytical performance report. OC-Sensor Ceres.; 2022.
88. Comparison IDK® TurbiFIT® and IDK® Hemoglobin ELISA. 2022.
89. Baeyens J-P, Serrien B, Goossens M, Clijsen R. Questioning the "SPIN and SNOUT" rule in clinical testing. *Archives of Physiotherapy* 2019;9:4.
90. Bailey J, Chapman C, Jones J, Oliver S, Morling J, Banerjea A*, et al.* Sociodemographic variations in the uptake of faecal immunochemical tests (fit) in a primary care symptomatic pathway for colorectal cancer. *Colorectal Disease* 2022;24(Supplement 3):8-9.
91. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996;313:275.
92. Westwood M, Corro Ramos I, Lang S, Luyendijk M, Zaim R, Stirk L*, et al.* Faecal immunochemical tests to triage patients with lower abdominal symptoms for suspected colorectal cancer referrals in primary care: a systematic review and cost-effectiveness analysis. *Health Technology Assessment (Winchester, England)* 2017;21:1-234.
93. Medina-Lara A, Grigore B, Lewis R, Peters J, Price S, Landa P*, et al.* Cancer diagnostic tools to aid decision-making in primary care: mixed-methods systematic reviews and cost-effectiveness analysis. *Health Technology Assessment (Winchester, England)* 2020;24:1-332.
94. National Institute for Health and Care Excellence. NICE health technology evaluations: the manual. Process and methods [PMG36]. London, UK: National Institute for Health and Care Excellence (NICE); 2022.
95. Neal RD, Tharmanathan P, France B, Din NU, Cotton S, Fallon-Ferguson J*, et al.* Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? Systematic review. *British Journal of Cancer* 2015;112:S92-S107.
96. Thomas C, Mandrik, O., Whyte, S. Development of the Microsimulation Model in Cancer of the Bowel (MiMiC-Bowel), an Individual Patient Simulation Model for Investigation of the Cost-effectiveness of Personalised Screening and Surveillance Strategies. *ScHARR HEDS Discussion Papers* 2020.
97. Bressler B, Paszat LF, Chen Z, Rothwell DM, Vinden C, Rabeneck L. Rates of New or Missed Colorectal Cancers After Colonoscopy and Their Risk Factors: A Population-Based Analysis. *Gastroenterology* 2007;132:96-102.
98. Atkin W, Dadswell E, Wooldrage K, Kralj-Hans I, von Wagner C, Edwards R*, et al.* Computed tomographic colonography versus colonoscopy for investigation of patients with symptoms suggestive of colorectal cancer (SIGGAR): a multicentre randomised trial. *The Lancet* 2013;381:1194-202.
99. Lin JS, Piper MA, Perdue LA, Rutter CM, Webber EM, O'Connor E*, et al.* Screening for Colorectal Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA* 2016;315:2576-94.
100. Martín-López JE, Beltrán-Calvo C, Rodríguez-López R, Molina-López T. Comparison of the accuracy of CT colonography and colonoscopy in the diagnosis of colorectal cancer. *Colorectal Disease* 2014;16:O82-O9.

101. Horsthuis K, Bipat S, Bennink RJ, Stoker J. Inflammatory Bowel Disease Diagnosed with US, MR, Scintigraphy, and CT: Meta-analysis of Prospective Studies. *Radiology* 2008;247:64-79.
102. Thomas C, Mandrik O, Whyte S. Development of the Microsimulation Model in Cancer of the Bowel (MiMiC-Bowel), an Individual Patient Simulation Model for Investigation of the Cost-effectiveness of Personalised Screening and Surveillance Strategies. 2020.
103. Kaltenthaler E, Tappenden P, Paisley S, Squires H. NICE DSU Technical Support Document 13: Identifying and reviewing evidence to inform the conceptualisation and population of cost-effectiveness models. In: Sheffield Uo, ed. DECISION SUPPORT UNIT. Sheffield, UK; 2011: 72.
104. Public Health England, Macmillan. UK complete cancer prevalence for 2013. Macmillan-NCRAS Cancer Prevalence Project. Updated UK Complete Cancer Prevalence for 2013 Workbook. In; 2017.
105. NHS Digital, National Cancer Registration and Analysis Service (NCRAS). Staging data in England. Stage group by Clinical Commissioning Group (CCG), Sustainability and Transformation Partnership (STP) or Cancer Alliance (CA) by cancer type for 21 cancer types 2019. In; 2021.
106. Pasvol T, Horsfall L, Bloom S, Segal A, Sabin C, Field N*, et al.* Incidence and prevalence of inflammatory bowel disease in UK primary care: a population-based cohort study. *BMJ Open* 2020;10:e036584.
107. Ghosh N, Premchand P. A UK cost of care model for inflammatory bowel disease. *Frontline Gastroenterology* 2015;6:169.
108. Lin JS, Perdue LA, Henrikson NB, Bean SI, Blasi PR. *Agency for Healthcare Research and Quality* 2021:05.
109. Gatto NM, Frucht H, Sundararajan V, Jacobson JS, Grann VR, Neugut AI. Risk of Perforation After Colonoscopy and Sigmoidoscopy: A Population-Based Study. *JNCI: Journal of the National Cancer Institute* 2003;95:230-6.
110. Office for National Statistics. National life tables: England 2018-2020. In; 2021.
111. National Institute for Health and Care Excellence. Upadacitinib for treating moderately to severely active ulcerative colitis [TA856]. Committee papers; 2022.
112. National institute for Health and Care Excellence. Vedolizumab for treating moderately to severely active ulcerative colitis [TA342]; 2015.
113. Stark RG, Reitmeir P, Leidl R, König H-H. Validity, reliability, and responsiveness of the EQ-5D in inflammatory bowel disease in Germany. *Inflammatory Bowel Diseases* 2010;16:42-51.
114. M Hernández Alava, S Pudney, A Wailoo. Estimating EQ-5D by age and sex for the UK. 2022.
115. Dorian P, Kongnakorn T, Phatak H, Rublee DA, Kuznik A, Lanitis T*, et al.* Cost-effectiveness of apixaban vs. current standard of care for stroke prevention in patients with atrial fibrillation. *European Heart Journal* 2014;35:1897-906.
116. Ara R, Brazier JE. Using Health State Utility Values from the General Population to Approximate Baselines in Decision Analytic Models when Condition-Specific Data are Not Available. *Value in Health* 2011;14:539-45.
117. Jones K, Burns A. Unit Costs of Health and Social Care: PSSRU; 2021.
118. NHS. National Cost Collection: National schedule of NHS costs - Year 2021/22 - NHS trusts and NHS foundation trusts. In; 2023.
119. Lyratzopoulos G, Abel GA, McPhail S, Neal RD, Rubin GP. Measures of promptness of cancer diagnosis in primary care: secondary analysis of national audit data on patients with 18 common and rarer cancers. 2013.
120. NHS Digital. Routes to Diagnosis, 2018. In: Digital N, ed.; 2022.
121. Jones KC, Weatherly H, Birch S, Castelli A, Chalkley M, Dargan A*, et al.* Unit Costs of Health and Social Care 2022 Manual. Technical report. Kent, UK; 2023.

122. Arasaradnam RP, Bhala N, Evans C, Greenaway J, Logan R, Penman I*, et al.* Faecal immunochemical testing in the COVID-19 era: balancing risk and costs. *The Lancet Gastroenterology & Hepatology* 2020;5:717-9.
123. Woehl A HA, McEwan P, editors. The relation between disease activity, quality of life and health utility in patients with ulcerative colitis. *Gut* 2008;57:A153.
124. Feagan BG, Rutgeerts P, Sands BE, Hanauer S, Colombel J-F, Sandborn WJ*, et al.* Vedolizumab as Induction and Maintenance Therapy for Ulcerative Colitis. *New England Journal of Medicine* 2013;369:699-710.
125. Moss S, Mathews C, Day TJ, Smith S, Seaman HE, Snowball J*, et al.* Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results from a pilot study within the national screening programme in England. *Gut* 2017;66:1631-44.
126. Stinnett AA, Mullahy J. Net Health Benefits: A New Framework for the Analysis of Uncertainty in Cost-Effectiveness Analysis. *Medical Decision Making* 1998;18:S68-S80.
127. Gies A, Gruner LF, Schrotz-King P, Brenner H. Effect of Imperfect Compliance With Instructions for Fecal Sample Collection on Diagnostic Performance of 9 Fecal Immunochemical Tests. *Clin Gastroenterol Hepatol* 2019;17:1829-39.e4.
128. James T, Nicholson BD, Marr R, Paddon M, East JE, Justice S*, et al.* Faecal immunochemical testing (FIT): sources of result variation based on three years of routine testing of symptomatic patients in English primary care. *British Journal of Biomedical Science* 2021;78:211-7.
129. Mellen S, de Ferrars M, Chapman C, Bevan S, Turvill J, Turnock D. Evaluation of sample stability for a quantitative faecal immunochemical test and comparison of two sample collection approaches. *Ann Clin Biochem* 2018;55:657-64.
130. O'Driscoll S, Carroll M, Maclean W, Piggott C, Jourdan I, Benton SC. Assessment of the analytical performance of point-of-care faecal immunochemical tests for haemoglobin. *Ann Clin Biochem* 2021;58:181-9.
131. O'Driscoll S, Piggott C, Bruce H, Benton SC. An evaluation of ten external quality assurance scheme (EQAS) materials for the faecal immunochemical test (FIT) for haemoglobin. 2021;59:307-13.
132. Piggott C, Carroll MRR, John C, O'Driscoll S, Benton SC. Analytical evaluation of four faecal immunochemistry tests for haemoglobin. *Clin Chem Lab Med* 2020;59:173-8.
133. Piggott C, Shugaa Z, Benton SC. Independent internal quality control (IQC) for faecal immunochemical tests (FIT) for haemoglobin: use of FIT manufacturers' IQC for other FIT systems. 2021;59:e41-e3.
134. Zubero MB, Arana-Arri E, Pijoan JI, Portillo I, Idigoras I, Lopez-Urrutia A*, et al.* Population-based colorectal cancer screening: comparison of two fecal occult blood test. *Frontiers in Pharmacology* 2014;4.
135. Ayling RM, Machesney M. Service evaluation of faecal immunochemical testing introduced for use in North East London for patients at low risk of colorectal cancer. *Journal of Clinical Pathology* 2021;74(3):163-6.
136. Bailey J, Ibrahim H, Bunce J, Chapman C, Morling J, Simpson J*, et al.* Quantitative FIT stratification is superior to NICE referral criteria NG12 in a Two-Week Wait Colorectal Cancer population. In. Submitted to IJCD 2020; 2020.
137. Bailey JA, Khawaja A, Andrews H, Weller J, Chapman C, Morling JR*, et al.* GP access to FIT increases the proportion of colorectal cancers detected on urgent pathways in symptomatic patients in Nottingham. *Surgeon* 2021b;19:93-102.
138. Bailey SER, Abel GA, Atkins A, Byford R, Davies S-J, Mays J*, et al.* Diagnostic performance of a faecal immunochemical test for patients with low-risk symptoms of colorectal cancer in primary care: an evaluation in the South West of England. *British Journal of Cancer* 2021;124:1231-6.

139. Chapman C, Bunce J, Oliver S, Ng O, Tangri A, Rogers R*, et al.* Service evaluation of faecal immunochemical testing and anaemia for risk stratification in the 2-week-wait pathway for colorectal cancer. *BJS Open* 2019;3(3):395-402.
140. Chapman C, Thomas C, Morling J, Tangri A, Oliver S, Simpson JA*, et al.* Early clinical outcomes of a rapid colorectal cancer diagnosis pathway using faecal immunochemical testing in Nottingham. *Colorectal Disease* 2020;22:679-88.
141. Ibrahim HAH, Chapman C, Morling J, Weller J, Tangri A, Simpson JA*, et al.* Keeping FIT: Faecal haemoglobin measurement with FIT has stratification value in the diagnosis of colorectal cancer in all symptom and age groups. *Colorectal Disease* 2019;21(Supplement 2):10.
142. Khasawneh F, Osborne R, Stephenson J, al. e. Faecal immunochemical testing is a cost-effective way to stratify symptomatic patients for urgent straight to test investigation. *Colorectal Dis* 2020;22:6-12.
143. McSorley ST, Digby J, Clyde D, Cruickshank N, Burton P, Barker L*, et al.* Yield of colorectal cancer at colonoscopy according to faecal haemoglobin concentration in symptomatic patients referred from primary care. *Colorectal Disease* 2021;23:1615-21.
144. Steele R, Fraser C. Haemoglobin for Timely Assessment of Patients with Symptoms of Colorectal Disease in Olsen Timely Diagnosis of Colorectal Disease. In: Springer; 2018.
145. Widlak MM, Neal M, Daulton E, Thomas CL, Tomkins C, Singh B*, et al.* Risk stratification of symptomatic patients suspected of colorectal cancer using faecal and urinary markers. *Colorectal Disease* 2018;20(12):O335-O42.
146. Widlak MM, Neal M, Daulton E, Thomas CL, Tomkins C, Singh B*, et al.* Risk stratification of symptomatic patients suspected of colorectal cancer using faecal and urinary markers. *Colorectal Dis* 2018;20:O335-o42.
147. Widlak MM, Thomas CL, Thomas MG, Tomkins C, Smith S, O'Connell N*, et al.* Diagnostic accuracy of faecal biomarkers in detecting colorectal cancer and adenoma in symptomatic patients. *Alimentary Pharmacology & Therapeutics* 2017;45:354-63.
148. Burke CA, Lieberman D, Feuerstein JD. AGA Clinical Practice Update on Approach to the Use of Noninvasive Colorectal Cancer Screening Options: Commentary. *Gastroenterology* 2022;162:952-6.
149. Craig M, Turner J, Torkington J, Crosby T. Faecal immunochemical test: challenges and opportunities for cancer diagnosis in primary care. *British Journal of General Practice* 2022;72:366-7.
150. D'Souza N, Anthony B, Muti A. Faecal immunochemical testing in general practice. *British Journal of General Practice* 2019;69:60.
151. Mowat C, Digby J, Strachan JA, Steele RJC, Fraser CG. Low Sensitivity of Fecal Immunochemical Tests (FIT) for Detection of Sessile Serrated Adenomas/Polyps Confirmed Over Clinical Setting, Geography, and FIT System. *Dig Dis Sci* 2019;64:3024-6.
152. Ray K. Prognostic potential of repeated faecal haemoglobin levels in CRC detection. *Nature Reviews Gastroenterology & Hepatology* 2022;19:416.
153. Rees CJ, Hamilton W. BSG guidelines on faecal immunochemical testing: Are they 'FIT' for purpose? *Gut* 2022;(no pagination).
154. Trivedi M, Gupta S. Is Promotion of Fecal Immunochemical Testing "FIT" to Address COVID-19 Disruptions to Colorectal Cancer Screening? *Gastroenterology* 2022;162:1761-2.
155. Ali O, Gupta S, Brain K, Lifford KJ, Paranjothy S, Dolwani S. Acceptability of alternative technologies compared with faecal immunochemical test and/or colonoscopy in colorectal cancer screening: A systematic review. *Journal of Medical Screening* 2022:9691413221109999.
156. Auge JM, Fraser CG, Rodriguez C, Roset A, Lopez-Ceron M, Grau J*, et al.* Clinical utility of one versus two faecal immunochemical test samples in the detection of

advanced colorectal neoplasia in symptomatic patients. *Clinical Chemistry & Laboratory Medicine* 2016;54:125-32.
157. Auge JM, Rodriguez C, Espanyol O, Rivero L, Sandalinas S, Grau J*, et al.* An evaluation of the SENTiFIT 270 analyser for quantitation of faecal haemoglobin in the investigation of patients with suspected colorectal cancer. *Clinical Chemistry & Laboratory Medicine* 2018;56:625-33.
158. Cahill C, Lipson ME, Afzal AR, Maclean AR, Wong CK, Roen S*, et al.* Improved Survival in a Cohort of Patients 75 years and over with FIT-Detected Colorectal Cancer. *Diseases of the Colon & Rectum* 2022;28:28.
159. Chandrapalan S, Hee SW, Widlak MM, Farrugia A, Alam MT, Smith S*, et al.* Performance of the faecal immunochemical test for the detection of colorectal neoplasms and the role of proton pump inhibitors in their diagnostic accuracy. *Colorectal Disease* 2021;23:1649-57.
160. Cubiella J, Vega P, Salve M, Díaz-Ondina M, Alves MT, Quintero E*, et al.* Development and external validation of a faecal immunochemical test-based prediction model for colorectal cancer detection in symptomatic patients. *BMC Medicine* 2016;14:128.
161. Digby J, Cleary S, Gray L, Datt P, Goudie DR, Steele RJC*, et al.* Faecal haemoglobin can define risk of colorectal neoplasia at surveillance colonoscopy in patients at increased risk of colorectal cancer. *United European Gastroenterology Journal* 2020;8:559-66.
162. Digby J, Strachan JA, McCann R, Steele RJC, Fraser CG, Mowat C. Measurement of faecal haemoglobin with a faecal immunochemical test can assist in defining which patients attending primary care with rectal bleeding require urgent referral. *Annals of Clinical Biochemistry* 2020;57:325-7.
163. Digby J, Strachan JA, Mowat C, Steele RJC, Fraser CG. Appraisal of the faecal haemoglobin, age and sex test (FAST) score in assessment of patients with lower bowel symptoms: an observational study. *BMC Gastroenterology* 2019;19:213.
164. Eskelinen M, Meklin J, Guimaraes DP, Selander T, Tiusanen T, Syrjanen K*, et al.* The ColonView (CV) Quick Test for Fecal Occult Blood Shows Significantly Higher Diagnostic Accuracy in Detecting Distal than Proximal Colorectal Cancer. *Anticancer Research* 2022;42:1879-91.
165. Gies A, Cuk K, Schrotz-King P, Brenner H. Direct comparison of ten quantitative fecal immunochemical tests for hemoglobin stability in colorectal cancer screening. *Clinical and Translational Gastroenterology* 2018;9.
166. Gies A, Cuk K, Schrotz-King P, Brenner H. Direct Comparison of Diagnostic Performance of 9 Quantitative Fecal Immunochemical Tests for Colorectal Cancer Screening. *Gastroenterology* 2018;154:93-104.
167. Gies A, Cuk K, Schrotz-King P, Brenner H. Combination of Different Fecal Immunochemical Tests in Colorectal Cancer Screening: Any Gain in Diagnostic Performance? *Journal* 2019.
168. Gies A, Niedermaier T, Alwers E, Hielscher T, Weigl K, Heisser T*, et al.* Consistent Major Differences in Sex- and Age-Specific Diagnostic Performance among Nine Faecal Immunochemical Tests Used for Colorectal Cancer Screening. *Cancers (Basel)* 2021;13.
169. Habbu PP, Ananthi N, Shaikh AK. Study of Specificity, Sensitivity, Efficiency & Clinical Correlation between Timp-1 and Mif Protein as Biochemical Markers in Colorectal Cancer Patients. *International Journal of Pharmaceutical Sciences and Research* 2022;13(8):3298-303.
170. Herrero J-M, Vega P, Salve M, Bujanda L, Cubiella J. Symptom or faecal immunochemical test based referral criteria for colorectal cancer detection in symptomatic patients: a diagnostic tests study. *BMC Gastroenterology* 2018;18:155.
171. Hicks G, D'Souza N, Georgiou Delisle T, Chen M, Benton SC, Abulafi M*, et al.* Using the faecal immunochemical test in patients with rectal bleeding: evidence from the NICE FIT study. *Colorectal Disease* 2021;23:1630-8.

172. Jin P, You P, Fang J, Kang Q, Gu F, Cai Y*, et al.* Comparison of Performance of Two Stool DNA Tests and a Fecal Immunochemical Test in Detecting Colorectal Neoplasm: A Multicenter Diagnostic Study. *Cancer Epidemiology, Biomarkers & Prevention* 2022;31:654-61.
173. Jin P, You P, Fang J, Kang Q, Gu F, Cai Y*, et al.* Comparison of Performance of Two Stool DNA Tests and a Fecal Immunochemical Test in Detecting Colorectal Neoplasm: A Multicenter Diagnostic Study. *Cancer Epidemiol Biomarkers Prev* 2022;31:654-61.
174. Kapidzic A, van Roon AH, van Leerdam ME, van Vuuren AJ, van Ballegooijen M, Lansdorp-Vogelaar I*, et al.* Attendance and diagnostic yield of repeated two-sample faecal immunochemical test screening for colorectal cancer. *Gut* 2017;66:118-23.
175. Kaul A, Shah A, Magill F, Hawkins S, Skaife P. Immunological faecal occult blood testing: a discriminatory test to identify colorectal cancer in symptomatic patients. *International Journal of Surgery* 2013;11:329-31.
176. Lincoln A, Benton S, Piggott C, North BV, Rigney J, Young C*, et al.* Exploring the utility and acceptability of Faecal immunochemical testing (FIT) as a novel intervention for the improvement of colorectal Cancer (CRC) surveillance in individuals with lynch syndrome (FIT for lynch study): a single-arm, prospective, multi-centre, non-randomised study. *BMC Cancer* 2022;22:1144.
177. Lu DC, Zhang QF, Li L, Luo XK, Liang B, Lu YH*, et al.* Methylated Septin9 has moderate diagnostic value in colorectal cancer detection in Chinese population: a multicenter study. *BMC Gastroenterology* 2022;22:232.
178. Lué A, Hijos G, Sostres C, Perales A, Navarro M, Barra MV*, et al.* The combination of quantitative faecal occult blood test and faecal calprotectin is a cost-effective strategy to avoid colonoscopies in symptomatic patients without relevant pathology. *Therap Adv Gastroenterol* 2020;13:1756284820920786.
179. Luthgens K, Maier A, Kampert I, Sieg A, Schmidt-Gayk H. Hemoglobin-haptoglobin-complex: a highly sensitive assay for the detection of fecal occult blood. *Clinical laboratory* 1998;44:543-51.
180. Mattar R, Marques SB, Minata MK, Silva-Etto J, Sakai P, EGH DEM. Diagnostic Accuracy of One Sample or Two Samples Quantitative Fecal Immunochemical Tests for Intestinal Neoplasia Detection. *Arquivos de Gastroenterologia* 2020;57:316-22.
181. Meester RGS, van de Schootbrugge-Vandermeer HJ, Breekveldt ECH, de Jonge L, Toes-Zoutendijk E, Kooyker A*, et al.* Faecal occult blood loss accurately predicts future detection of colorectal cancer. A prognostic model. *Gut* 2022;10:10.
182. Navarro M, Hijos G, Sostres C, Lue A, Puente-Lanzarote JJ, Carrera-Lasfuentes P*, et al.* Reducing the Cut-Off Value of the Fecal Immunochemical Test for Symptomatic Patients Does Not Improve Diagnostic Performance. *Frontiers in Medicine* 2020;7:410.
183. Navarro M, Hijos G, Sostres C, Lué A, Puente-Lanzarote JJ, Carrera-Lasfuentes P*, et al.* Reducing the Cut-Off Value of the Fecal Immunochemical Test for Symptomatic Patients Does Not Improve Diagnostic Performance. *Frontiers in Medicine* 2020;7.
184. Parente F, Marino B, Ilardo A, Fracasso P, Zullo A, Hassan C*, et al.* A combination of faecal tests for the detection of colon cancer: a new strategy for an appropriate selection of referrals to colonoscopy? A prospective multicentre Italian study. *Eur J Gastroenterol Hepatol* 2012;24:1145-52.
185. Rodríguez-Alonso L, Rodríguez-Moranta F, Ruiz-Cerulla A, Lobatón T, Arajol C, Binefa G*, et al.* An urgent referral strategy for symptomatic patients with suspected colorectal cancer based on a quantitative immunochemical faecal occult blood test. *Digestive and Liver Disease* 2015;47:797-804.
186. Small S, Coulson R, Spence R, McAllister I. Is qFIT a useful tool in prioritising symptomatic patients referred with suspect colorectal cancer in the COVID-19 era? *Ulster Medical Journal* 2022;91:79-84.
187. Suehiro Y, Zhang Y, Hashimoto S, Takami T, Higaki S, Shindo Y*, et al.* Highly sensitive faecal DNA testing of TWIST1 methylation in combination with faecal

immunochemical test for haemoglobin is a promising marker for detection of colorectal neoplasia. *Annals of Clinical Biochemistry* 2018;55(1):59-68.
188. Switalski J, Tatara T, Wnuk K, Miazga W, Karauda D, Matera A*, et al.* Clinical Effectiveness of Faecal Immunochemical Test in the Early Detection of Colorectal Cancer-An Umbrella Review. *Cancers* 2022;14:09.
189. Tsapournas G, Hellstrom PM, Cao Y, Olsson LI. Diagnostic accuracy of a quantitative faecal immunochemical test vs. symptoms suspected for colorectal cancer in patients referred for colonoscopy. *Scandinavian Journal of Gastroenterology* 2020;55:184-92.
190. van Turenhout ST, Oort FA, van der Hulst RWM, Visscher AP, Terhaar sive Droste JS, Scholten P*, et al.* Prospective cross-sectional study on faecal immunochemical tests: sex specific cut-off values to obtain equal sensitivity for colorectal cancer? *BMC Gastroenterology* 2014;14:217.
191. Xu H, Chen H, Hu J, Xiong Z, Li D, Wang S*, et al.* Feasibility of quantification based on novel evaluation with stool DNA and fecal immunochemical test for colorectal cancer detection. *BMC Gastroenterology* 2022a;22:384.
192. Xu J, Rong L, Gu F, You P, Ding H, Zhai H*, et al.* Asia-Pacific Colorectal Screening Score Combined With Stool DNA Test Improves the Detection Rate for Colorectal Advanced Neoplasms. *Clinical Gastroenterology & Hepatology* 2022b;14:14.
193. Young GP, Woodman RJ, Symonds E. Detection of advanced colorectal neoplasia and relative colonoscopy workloads using quantitative faecal immunochemical tests: an observational study exploring the effects of simultaneous adjustment of both sample number and test positivity threshold. *BMJ Open Gastroenterology* 2020;7:09.
194. Zhao S, Wang S, Pan P, Xia T, Wang R, Cai Q*, et al.* FIT-based risk-stratification model effectively screens colorectal neoplasia and early-onset colorectal cancer in Chinese population: a nationwide multicenter prospective study. *Journal of hematology & oncology* 2022;15:162.
195. Calanzani N, Pannebakker MM, Tagg MJ, Walford H, Holloway P, de Wit N*, et al.* Who are the patients being offered the faecal immunochemical test in routine English general practice, and for what symptoms? A prospective descriptive study. *BMJ Open* 2022;12:e066051.
196. Carroll MRR, John C, Mantio D, Djedovic NK, Benton SC. An assessment of the effect of haemoglobin variants on detection by faecal immunochemical tests. *Annals of Clinical Biochemistry* 2018;55(6):706-9.
197. Chen CC, Chang PY, Chang YS, You JF, Chan EC, Chen JS*, et al.* MicroRNA-based signature for diagnosis and prognosis of colorectal cancer using residuum of fecal immunochemical test. *Biomedical Journal* 2022;22:22.
198. Fernandez de Castro JD, Baiocchi Ureta F, Fernandez Gonzalez R, Pin Vieito N, Cubiella Fernandez J. Faecal Immunochemical Test Impact on Prognosis of Colorectal Cancer Detected in Symptomatic Patients. *Diagnostics* 2022;12:17.
199. Hunt N, Allcock R, Myers M. Faecal immunochemical testing (FIT) for colorectal cancer in symptomatic primary care patients. Paper presented at: Clinica Chimica Acta.
200. Lee J-M, Park MJ, Heo W, Park KG, Park YG, Han SB*, et al.* Clinical Utility of Fecal Immunochemical Transferrin Test in Gastrointestinal Bleeding Detection. *acm* 2018;21:51-7.
201. Maria Theresa R, Matthew R, Richard MM, Fareeda C, Mona J, Julia W. Rapid diagnostic pathways for suspected colorectal cancer: views of primary and secondary care clinicians on challenges and their potential solutions. *BMJ Open* 2015;5:e008577.
202. Navarro M, Hijos G, Ramirez T, Omella I, Carrera-Lasfuentes P, Lanas Á. Fecal Hemoglobin Concentration, a Good Predictor of Risk of Advanced Colorectal Neoplasia in Symptomatic and Asymptomatic Patients. *Front Med (Lausanne)* 2019;6:91.

203. Niedermaier T, Alwers E, Chen X, Heisser T, Hoffmeister M, Brenner H. A single measurement of fecal hemoglobin concentration outperforms polygenic risk score in colorectal cancer risk assessment. *medRxiv* 2022;22.
204. Wilson N, Baker-Beal L, Kyaw WW, Sonabend R, Logan R. Real world experience of faecal immunochemical testing (FIT) in uk primary care to support the referral and diagnosis of colorectal cancer (CRC). *United European Gastroenterology Journal* 2020;8(8 SUPPL):579-80.
205. Krivec. ASSESSMENT OF THE DIAGNOSTIC APPLICABILITY OF QUANTITATIVE IMMUNOCHEMICAL FAECAL OCCULT BLOOD TESTS. 2011.
206. Sandhu K, Naik S, Ayling RM. Use of faecal immunochemical testing as an alternative to faecal calprotectin in children. *Ann Clin Biochem* 2021;58:230-5.
207. Zacharopoulou L, Cama R, Kapoor N, Mebarek L, Bhatti H, Sawyer P*, et al.* PTH-99 Faecal Immunochemical Testsfor younger patients presenting with bowel symptoms. *Gut* 2021;70:A162-A.
208. Zhu M, Fan L, Han M, Zhu S, Zhang S, Shi H. The usefulness of fecal hemoglobin and calprotectin tests in diagnosing significant bowel diseases: a prospective study. *Scandinavian Journal of Gastroenterology* 2022:1-7.
209. Rodriguez-Alonso L, Rodriguez-Moranta F, Maisterra S, Botargues JM, Berrozpe A, Ruiz-Cerulla A*, et al.* The EPAGE guidelines are not an effective strategy for managing colonoscopies during the COVID-19 pandemic. *Gastroenterologia y Hepatologia* 2022;45:9-17.
210. Almoneef NM, Alkhenizan AH, Mahmoud AS, Alsoghayer SA, Aldheshe AA. The yield of fecal occult blood testing as a screening tool for colon cancer in a primary care setting. *Journal of Family Medicine & Primary Care* 2022;11:4435-9.
211. Cubiella J, Digby J, Rodríguez-Alonso L, Vega P, Salve M, Díaz-Ondina M*, et al.* The fecal hemoglobin concentration, age and sex test score: Development and external validation of a simple prediction tool for colorectal cancer detection in symptomatic patients. *Int J Cancer* 2017;140:2201-11.
212. Khan AA, Klimovskij M, Harshen R. Accuracy of faecal immunochemical testing in patients with symptomatic colorectal cancer. *Bjs Open* 2020;4.
213. Maclean W, Benton SC, Whyte MB, Rockall T, Jourdan I. Efficacy and accuracy of faecal sampling by a digital rectal examination for FIT. *Annals of Clinical Biochemistry* 2023:00045632231155021.
214. McDonald P, Digby J, Innes C, Strachan J, Carey F, Steele R*, et al.* Low faecal haemoglobin concentration potentially rules out significant colorectal disease. *Colorectal Disease* 2013;15:e151-e9.
215. County RÖ, Härjedalen RJ, Östergötland R, Uppsala County Council S, Dalarna County Council S, Värmland Li. Accuracy and Predictive Values for Colorectal Cancer of Quantitative FIT in Symptomatic Patients in Primary Care. In: https://ClinicalTrials.gov/show/NCT05156307; 2021.
216. Ghada Mohamed ABJS. Sensitivity and specificity of stool markers in Inflammatory bowel disease: A systematic review and meta-analysis.
217. Jen-Hao Yeh W-LW. Efficacy of fecal immunochemical test in young patients (aged ≤ 50 years).
218. Jennifer Pham ESMWJWGL-L. The diagnostic accuracy of faecal immunochemical tests for detecting colorectal cancer and pre-cancerous neoplasia in patients with iron deficiency with or without anaemia.
219. Loov A, Hogberg C, Lilja M, Theodorsson E, Hellstrom P, Metsini A*, et al.* Diagnostic accuracy for colorectal cancer of a quantitative faecal immunochemical test in symptomatic primary care patients: a study protocol. *Diagnostic and Prognostic Research* 2022;6:16.
220. Tian Zhi Lim JLAUJLGW. Uncovering the barriers and facilitators towards undergoing follow-up colonoscopy among individuals with FIT positive results: A systematic review.

221. Tian Zhi Lim JLJLGW. What can we do to improve compliance to follow-up colonoscopy among individuals with FIT positive results? A systematic review.
222. University M. FIT and Fecal Calprotectin in Patients With Chronic Lower GI Symptoms. In: https://ClinicalTrials.gov/show/NCT05514561; 2020.
223. Winnie Poulsen MTETWX. Factors associated with the quality of faecal occult blood test (FOBT) for early detection of colorectal cancer: a systematic review and meta-analysis protocol.
224. Zhen Junhai CB. Value of Faecal immunochemical tests in significant bowel disease screening among patients with lower bowel symptoms in primary care：a meta analysis.
225. Farkas NG, Fraser CG, Maclean W, Jourdan I, Rockall T, Benton SC. Replicate and repeat faecal immunochemical tests in symptomatic patients: A systematic review. *Annals of Clinical Biochemistry* 2022:45632221096036.
226. Jung YS, Im E, Park CH. Impact of antiplatelet agents and anticoagulants on the performance of fecal immunochemical tests: a systematic review and meta-analysis. *Surgical Endoscopy* 2022;36:4299-311.
227. Nasir Kansestani A, Zare ME, Tong Q, Zhang J. Comparison of faecal protein biomarkers' diagnostic accuracy for colorectal advanced neoplasms: a systematic review and meta-analysis. *Scientific Reports* 2022;12:2623.
228. Pang SJ, Lin ZP, Sun Z, Zhang Y, Yuan ZG, Yang N. Impact of antithrombotic drugs on the accuracy of fecal occult blood testing for advanced colorectal neoplasia screening: a meta-analysis and systematic review. *Zeitschrift fur Gastroenterologie* 2022;17:17.
229. Pin Vieito N, Zarraquinos S, Cubiella J. High-risk symptoms and quantitative faecal immunochemical test accuracy: Systematic review and meta-analysis. *World Journal of Gastroenterology* 2019;25:2383-401.
230. Pin-Vieito N, Tejido-Sandoval C, de Vicente-Bielza N, Sanchez-Gomez C, Cubiella J. Faecal immunochemical tests safely enhance rational use of resources during the assessment of suspected symptomatic colorectal cancer in primary care: systematic review and meta-analysis. *Gut* 2022;71:950-60.
231. Quyn AJ, Steele RJC, Digby J, Strachan JA, Mowat C, McDonald PJ*, et al.* Application of NICE guideline NG12 to the initial assessment of patients with lower gastrointestinal symptoms: not FIT for purpose? *Annals of Clinical Biochemistry* 2018;55(1):69-76.
232. Saw KS, Liu C, Xu W, Varghese C, Parry S, Bissett I. Faecal immunochemical test to triage patients with possible colorectal cancer symptoms: meta-analysis. *British Journal of Surgery* 2022;109:182-90.
233. Westwood M, Lang S, Armstrong N, van Turenhout S, Cubiella J, Stirk L*, et al.* Faecal immunochemical tests (FIT) can help to rule out colorectal cancer in patients presenting in primary care with lower abdominal symptoms: a systematic review conducted to inform new NICE DG30 diagnostic guidance. *BMC Medicine* 2017;15:189.
234. Zou J, Xiao Z, Wu Y, Yang J, Cui N. Noninvasive fecal testing for colorectal cancer. *Clinica Chimica Acta* 2022;524:123-31.
235. Cilona A, Zullo A, Hassan C, Ridola L, Annese M. Is faecal-immunochemical test useful in patients with iron deficiency anaemia and without overt bleeding? *Digestive and Liver Disease* 2011;43:1022-4.
236. Díaz O, Blanco V, Ceballos O, Salve M, Macía P, Cubiella J. CLINICAL OR ANALYTICAL CRITERIA FOR COLORECTAL CANCER (CRC) DETECTION IN SYMPTOMATIC PATIENTS? A DIAGNOSTIC TESTS STUDY. Paper presented at: IFCC WorldLab; Istanbul.
237. Berger BM, Schroy PC, Dinh TA. Screening for Colorectal Cancer Using a Multitarget Stool DNA Test: Modeling the Effect of the Intertest Interval on Clinical Effectiveness. *Clinical Colorectal Cancer* 2016;15:e65-e74.

238. Farkas N, O'Brien JW, Palyvos L, Maclean W, Benton S, Rockall T*, et al.* The increasing burden of the 2-week wait colorectal cancer pathway in a single centre: the impact of faecal immunochemical tests. *Annals of the Royal College of Surgeons of England* 2023:rcsann20220138.
239. Gendia A, Rottenburg H, Tam A, Faux W. Assessing quantitative FIT within the colorectal two-week wait pathway as a triage tool to focus endoscopy services. *British Journal of Surgery* 2021;108(SUPPL 7):vii7.
240. Hijos G, Lue A, Sostres C, Barra MV, Mascialino B, Andalucia C*, et al.* Combination of Quantitative Faecal Occult Blood Test and Fecal Calprotectin is a Cost-Effective Strategy to Avoid Non Pathological Colonoscopies in Symptomatic Patients. *Gastroenterology* 2018;154(6 Supplement 1):S-450-S-1.
241. Kearsey CC, Graham C, Lobb HS, Chacko J, Weatherburn R, Rooney PS. Cost effectiveness of using Faecal Immunochemical Testing (FIT) as an initial diagnostic investigation for patients with lower gastrointestinal symptoms suggestive of malignancy. *BMC Family Practice* 2021;22:90.
242. Kearsey C, Graham C, Chako J, Weatherburn R, Kennedy S, Rooney P. Cost comparison of faecal immunochemical tests to conventional methods as diagnostic tool in colorectal cancer. *Gut* 2021;70(SUPPL 1):A197-A8.
243. Khasawneh F, Osborne T, Danaher P, Barnes D, Chapman CJ, Stephenson JA*, et al.* Faecal immunochemical testing reduces demand and improves yield of Leicester's 2-week pathway for change in bowel habit. *Colorectal Disease* 2022;07:07.
244. Law R, Das A, Gregory D, Komanduri S, Muthusamy R, Rastogi A*, et al.* Endoscopic resection is cost-effective compared with laparoscopic resection in the management of complex colon polyps: an economic analysis. *Gastrointestinal Endoscopy* 2016;83:1248-57.
245. Lobo A, Torrejon Torres R, McAlindon M, Panter S, Leonard C, van Lent N*, et al.* Economic analysis of the adoption of capsule endoscopy within the British NHS. *International Journal for Quality in Health Care* 2020;32:332-41.
246. Lue A, Hijos G, Sostres C, Perales A, Navarro M, Barra MV*, et al.* The combination of quantitative faecal occult blood test and faecal calprotectin is a cost-effective strategy to avoid colonoscopies in symptomatic patients without relevant pathology. *Therapeutic Advances in Gastroenterology* 2020;13.
247. Padula WV, Millis MA, Worku AD, Pronovost PJ, Bridges JFP, Meltzer DO. Individualized cost-effectiveness analysis of patient-centered care: a case series of hospitalized patient preferences departing from practice-based guidelines. *Journal of Medical Economics* 2017;20(3):288-96.
248. Petersen MM, Ferm L, Kleif J, Piper TB, Romer E, Christensen IJ*, et al.* Triage may improve selection to colonoscopy and reduce the number of unnecessary colonoscopies. *Cancers* 2020;12(9):1-9.
249. Rodriguez-Alonso L, Rodriguez-Moranta F, Ruiz-Cerulla A, Arajol C, Serra K, Gilabert P*, et al.* The use of faecal immunochemical testing in the decision-making process for the endoscopic investigation of iron deficiency anaemia. *Clinical Chemistry & Laboratory Medicine* 2020;58:232-9.
250. Saing S, Haywood P, Duncan JK, Ma N, Cameron AL, Goodall S. Cost-effective imaging for resectability of liver lesions in colorectal cancer: an economic decision model. *ANZ Journal of Surgery* 2018;88:E507-E11.
251. Smith DH, O'Keeffe Rosetti M, Mosen DM, Rosales AG, Keast E, Perrin N*, et al.* Balancing Adherence and Expense: The Cost-Effectiveness of Two-Sample vs One-Sample Fecal Immunochemical Test. *Population Health Management* 2019;22:83-9.
252. Yu J, Ho V, Hoang T, Dolph M, Kwon CS, Forsythe A*, et al.* Opportunities To Enhance The Utility Of Economic Modeling In Colorectal Cancer: A Systematic Literature Review. *Gastroenterology* 2021;160(6 Supplement):S-603.
253. Brar HS, Kang J, Bathina P, Paine E. Trends and cost burden analysis of inpatient fecal occult blood testing in upper GI bleed. *Journal of Investigative Medicine* 2022;70(2):665.

254. Braun AL, Kassner A, Syrogiannouli L, Selby K, Bulliard JL, Martin Y*, et al.* Association between colorectal cancer testing and insurance type: Evidence from the Swiss Health Interview Survey 2012. *Preventive Medicine Reports* 2020;19:101111.
255. Coury J, Ramsey K, Gunn R, Judkins J, Davis M. Source matters: a survey of cost variation for fecal immunochemical tests in primary care. *BMC Health Services Research* 2022;22:204.
256. Fisher DA, Engel-Nitz N, Miller-Wilson LA, Le L, Limburg PJ. Patterns of colorectal cancer (CRC) screening rates among the average risk U.S. population. *Journal of Clinical Oncology Conference* 2022;40.
257. Fisher DA, Princic N, Miller-Wilson LA, Wilson K, Limburg P. Costs of colorectal cancer screening with colonoscopy, including post-endoscopy events, among adults with Medicaid insurance. *Current Medical Research & Opinion* 2022;38:793-801.
258. Paszat L, Sutradhar R, Luo J, Rabeneck L, Tinmouth J, Baxter NN. Overall Health Care Cost During the Year Following Diagnosis of Colorectal Cancer Stratified by History of Colorectal Evaluative Procedures. *Journal of the Canadian Association of Gastroenterology* 2021;4:274-83.
259. Pelitari S, Gautham A, Mistry P, Mohan S, Brookes M, McKaig B*, et al.* Impact on healthcare resources of switch from fecal occult blood test to fecal immunochemical test within the English Bowel Cancer Screening Program: a single-center study. *Gastrointestinal Endoscopy* 2021;94:598-606.
260. Subramanian S, Tangka FKL, Hoover S, Cole-Beebe M, Joseph D, DeGroff A. Comparison of Program Resources Required for Colonoscopy and Fecal Screening: Findings From 5 Years of the Colorectal Cancer Control Program. *Preventing Chronic Disease* 2019;16:E50.
261. Veettil SK, Kew ST, Lim KG, Phisalprapa P, Kumar S, Lee YY*, et al.* Very-low-dose aspirin and surveillance colonoscopy is cost-effective in secondary prevention of colorectal cancer in individuals with advanced adenomas: network meta-analysis and cost-effectiveness analysis. *BMC Gastroenterology* 2021;21:130.
262. Bobridge A, Bampton P, Cole S, Lewis H, Young G. The psychological impact of participating in colorectal cancer screening by faecal immuno-chemical testing--the Australian experience. *British Journal of Cancer* 2014;111:970-5.
263. Bobridge A, Young GP, Cole SR, Bampton PA. Does participating in a national bowel cancer screening program have a psychological impact? *Gastroenterology* 2012;1):S399.
264. Ferrari S, Mancini S, Alboni S, Artoni C, Fabbrizzi A, Feltri L*, et al.* Role of metabolic, atherogenetic and psychological factors in patients with colorectal adenomas: Preliminary results of the psycho-Neuro-Endocrino-Immunology Modena (PNEI-MO) Research Group. *European Psychiatry* 2016;33(SUPPL.):S79.
265. Kapidzic A, Korfage IJ, van Dam L, van Roon AH, Reijerink JC, Zauber AG*, et al.* Quality of life in participants of a CRC screening program. *British Journal of Cancer* 2012;107:1295-301.
266. Kirkoen B, Berstad P, Botteri E, Avitsland TL, Ossum AM, de Lange T*, et al.* Do no harm: no psychological harm from colorectal cancer screening. *British Journal of Cancer* 2016;114:497-504.
267. Kirkoen B, Berstad P, Botteri E, Bernklev L, El-Safadi B, Hoff G*, et al.* Psychological effects of colorectal cancer screening: Participants vs individuals not invited. *World Journal of Gastroenterology* 2016;22:9631-41.
268. Kirkoen B, Berstad P, De Lange T, Hoff G, Bernklev T. Psychological effects of colorectal cancer screening invitations: A randomized trial. *United European Gastroenterology Journal* 2014;1):A231.
269. Mahabaleshwarkar R, Khanna R, West-Strum D, Yang Y. Association between health-related quality of life and colorectal cancer screening. *Population Health Management* 2013;16:178-89.
270. Marshall DA, Johnson FR, Kulin NA, Ozdemir S, Walshd JME, Marshall JK*, et al.* How do physician assessments of patient preferences for colorectal cancer screening

tests differ from actual preferences? A comparison in Canada and the United States using a stated-choice survey. *Health Economics* 2009;18(12):1420-39.
271. Miles A, McClements PL, Steele RJ, Redeker C, Sevdalis N, Wardle J. The Psychological Impact of a Colorectal Cancer Diagnosis Following a Negative Fecal Occult Blood Test Result. *Cancer Epidemiology, Biomarkers & Prevention* 2015;24:1032-8.
272. Mountifield RE, Bampton PA, Prosser R, Bobridge A, Mikocka-Walus AA, Andrews JM. FIT+ and IBD individuals have differing psychological reactions to the need for colonoscopy. *Gastroenterology* 2013;1):S705.
273. Van Dam L, Van Leerdam ME, Hol L, Van Roon AH, Reijerink JC, Van Ballegooijen M*, et al.* Comparison of participants and non-participants in a flexible sigmoidoscopy screening program, with an alternative invitation for fecal immunochemical testing. *Gastroenterology* 2010;1):S351.
274. Vermeer NCA, van der Valk MJM, Snijders HS, Vasen HFA, Gerritsen van der Hoop A, Guicherit OR*, et al.* Psychological distress and quality of life following positive fecal occult blood testing in colorectal cancer screening. *Psycho-Oncology* 2020;29:1084-91.
275. Wattchow DA, Weller DP, Esterman A, Pilotto LS, McGorm K, Hammett Z*, et al.* General practice vs surgical-based follow-up for patients with colon cancer: randomised controlled trial. *British Journal of Cancer* 2006;94:1116-21.
276. Whynes DK, Neilson AR, Robinson MH, Hardcastle JD. Colorectal cancer screening and quality of life. *Quality of Life Research* 1994;3:191-8.
277. Yamada E, Watanabe S, Nakajima A. Associations of Mental Health and Physical Function with Colonoscopy-related Pain. *Internal Medicine* 2017;56:383-8.
278. England N. Waiting Times for Suspected and Diagnosed Cancer Patients: 2020-21 Annual Report; 2021.
279. England N. Waiting Times for Suspected and Diagnosed Cancer Patients: 2019-20 Annual Report; 2020.
280. Service NCRaA. Cancer Waiting Times (CWT) urgent suspected cancer referrals: referral, conversion and detection rates. In: CancerData, ed.; 2019.
281. Service NCRaA. Staging Data In England. In: CancerData, ed.; 2019.
282. Jones E, Epstein D, Garcia-Mochon L. A Procedure for Deriving Formulas to Convert Transition Rates to Probabilities for Multistate Markov Models. *Med Decis Making* 2017;37:779-89.
283. Statistics OfN. National Life Tables: England. In; 2021.
284. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. *Value Health* 2010;13:509-18.
285. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence Synthesis for Decision Making 2: A Generalized Linear Modeling Framework for Pairwise and Network Meta-analysis of Randomized Controlled Trials. *Medical Decision Making* 2012;33:607-17.
286. Sturtz S, Ligges U, Gelman A. R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software* 2005;12:1 - 16.

8 APPENDICES

Appendix 1: Literature search strategies

A. Clinical review search strategy

Ovid MEDLINE(R) ALL <1946 to December 06, 2022>

1 f?ecal immunochemical test.mp. 1259
2 f?ecal occult blood.mp. 4447
3 f?ecal h?emoglobin.mp. 269
4 ((immunochromatographic or immuno-chromatographic or immunochem$ or immuno-chem$ or immunohistochem$ or immuno-histochem$ or immunol$ or immunoassay or immuno* assay or immunoturbidimetric or immunosorbent or elisa) adj4 (f?ecal or f?eces or stool or stools or FIT)).mp. 3598
5 (iFOBT or qFIT).mp. 208
6 or/1-5 7290
7 F?ecal h?emoglobin.ti,ab,ot,hw. 256
8 H?emoccult.ti,ab,ot,hw. 728
9 FOBT.ti,ab,ot,hw. 1429
10 7 or 8 or 9 2335
11 (f?ecal or f?eces or stool or stools).ti,ab,ot,hw. 211912
12 occult blood/ or occult blood.ti,ab,ot,hw. 8924
13 (test$ or measur$ or screen$ or exam$).ti,ab,ot,hw. 10705030
14 11 and 12 and 13 6023
15 6 or 10 or 14 8736
16 exp colorectal neoplasms/ 231240
17 exp cecal neoplasms/ 6041
18 ((colorect$ or rectal$ or rectum$ or colon$ or sigma$ or sigmo$ or rectosigm$ or bowel$ or anal or anus) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot,hw. 324030
19 CRC.ti,ab,ot. 43421
20 ((cecum or cecal or caecum or caecal or il?eoc?ecal or il?eoc?ecum) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 2755
21 (large intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 1839
22 (lower intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 34
23 16 or 17 or 18 or 19 or 20 or 21 or 22335879
24 15 and 23 5937
25 limit 24 to yr="2022 -Current" 426
26 (FOB gold$ or FOBgold$ or SENTiFIT).mp. 38
27 (JACK-arc$ or JACKarc$ or HM-JACK$ or HM JACK$ or HMJACK$).mp. 23
28 (OC Sensor$ or OC-Sensor$ or OCSensor$ or Ceres).mp. 371
29 (OC Pledia$ or OC-Pledia$ or OCPledia or OC-iO).mp. 0
30 (NS-Prime or NSPrime or NS-Plus).mp. 37

31 (POC FIT QRG or POCFITQRG).mp. 0
32 (immundiagnostik or IDK or turbifit or turbitube).mp. 125
33 quikread.mp. 19
34 or/25-33 994
35 limit 34 to yr="2016 -Current" 740
36 exp animals/ not (exp animals/ and humans/)5072762
37 35 not 36 729

Embase <1974 to 2022 Week 49> searched 6th December 2022

1 f?ecal immunochemical test.mp. 2253
2 f?ecal occult blood.mp. 6908
3 f?ecal h?emoglobin.mp. 436
4 ((immunochromatographic or immuno-chromatographic or immunochem$ or immuno-chem$ or immunohistochem$ or immuno-histochem$ or immunol$ or immunoassay or immuno* assay or immunoturbidimetric or immunosorbent or elisa) adj4 (f?ecal or f?eces or stool or stools or FIT)).mp. 6067
5 (iFOBT or qFIT).mp. 392
6 or/1-5 11749
7 F?ecal h?emoglobin.ti,ab,ot,hw. 422
8 H?emoccult.ti,ab,ot,hw. 987
9 FOBT.ti,ab,ot,hw. 2786
10 7 or 8 or 9 4077
11 (f?ecal or f?eces or stool or stools).ti,ab,ot,hw. 279057
12 occult blood/ or occult blood.ti,ab,ot,hw. 18102
13 (test$ or measur$ or screen$ or exam$).ti,ab,ot,hw. 13879472
14 11 and 12 and 13 10328
15 6 or 10 or 14 14766
16 exp colorectal cancer/ or colon cancer/ or rectum cancer/ 316446
17 exp cecum tumor/ 2471
18 ((colorect$ or rectal$ or rectum$ or colon$ or sigma$ or sigmo$ or rectosigm$ or bowel$ or anal or anus) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot,hw. 506633
19 CRC.ti,ab,ot. 70056
20 ((cecum or cecal or caecum or caecal or il?eoc?ecal or il?eoc?ecum) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 3559
21 (large intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 1807
22 (lower intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 44
23 16 or 17 or 18 or 19 or 20 or 21 or 22515018
24 15 and 23 9914
25 limit 24 to yr="2022 -Current" 649
26 (FOB gold$ or FOBgold$ or SENTiFIT).mp. 107
27 (JACK-arc$ or JACKarc$ or HM-JACK$ or HM JACK$ or HMJACK$).mp. 73

28 (OC Sensor$ or OC-Sensor$ or OCSensor$ or Ceres).mp. 774
29 (OC Pledia$ or OC-Pledia$ or OCPledia or OC-iO).mp. 0
30 (NS-Prime or NSPrime or NS-Plus).mp. 75
31 (POC FIT QRG or POCFITQRG).mp. 0
32 (immundiagnostik or IDK or turbifit or turbitube).mp. 411
33 quikread.mp. 52
34 or/25-33 1997
35 limit 34 to yr="2016 -Current" 1406
36 limit 35 to embase 732
37 limit 35 to conference abstracts 500
38 limit 35 to "preprints (unpublished, non-peer reviewed)" 7

The Cochrane Library (searched 12th December 2022)

Search Name: DAP50 final
Date Run: 12/12/2022 18:29:15
ID Search Hits
#1 (fecal immunochemical test* or faecal immunochemical test*):ti,ab,kw (Word variations have been searched) 497
#2 (fecal occult blood or faecal occult blood):ti,ab,kw (Word variations have been searched) 1087
#3 (fecal hemoglobin or faecal hemoglobin or fecal haemoglobin or faecal haemoglobin):ti,ab,kw (Word variations have been searched) 298
#4 ((immunochromatographic or immuno-chromatographic or immunochem* or immuno-chem* or immunohistochem* or immuno-histochem* or immunol* or immunoassay or immuno* assay or immunoturbidimetric or immunosorbent or elisa) near/4 (fecal or faecal or feces or faeces or stool or stools or FIT)):ti,ab,kw (Word variations have been searched) 1041
#5 (iFOBT or qFIT):ti,ab,kw (Word variations have been searched) 37
#6 (Hemoccult or haemoccult):ti,ab,kw (Word variations have been searched) 129
#7 (FOBT):ti,ab,kw (Word variations have been searched) 411
#8 ((fecal or feces or faecal or faeces or stool or stools)):ti,ab,kw AND (occult blood):ti,ab,kw AND (test* or measur* or screen* or exam*):ti,ab,kw (Word variations have been searched) 1153
#9 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 2191
#10 MeSH descriptor: [Colorectal Neoplasms] explode all trees 9373
#11 MeSH descriptor: [Cecal Neoplasms] explode all trees 21
#12 ((colorect* or rectal* or rectum* or colon* or sigma* or sigmo* or rectosigm* or bowel* or anal or anus) near/3 (cancer* or neoplas* or oncolog* or malignan* or tumo* or carcinoma* or adenocarcinoma* or sarcoma* or adenom* or lesion*)):ti,ab,kw (Word variations have been searched) 26503
#13 (CRC):ti,ab,kw (Word variations have been searched) 5111
#14 ((cecum or cecal or caecum or caecal or ileocecal or ileocecum or ileocaecal or ileocaecum) near/3 (cancer* or neoplas* or oncolog* or malignan* or tumo* or carcinoma* or adenocarcinoma* or sarcoma* or adenom* or lesion*)):ti,ab,kw (Word variations have been searched) 246
#15 (large intestin* near/3 (cancer* or neoplas* or oncolog* or malignan* or tumo* or carcinoma* or adenocarcinoma* or sarcoma* or adenom* or lesion*)):ti,ab,kw (Word variations have been searched) 172

#16 (lower intestin* near/3 (cancer* or neoplas* or oncolog* or malignan* or tumo* or carcinoma* or adenocarcinoma* or sarcoma* or adenom* or lesion*)):ti,ab,kw (Word variations have been searched) 182
#17 #10 or #11 or #12 or #13 or #14 or #15 or #16 27348
#18 #9 and #17 with Cochrane Library publication date Between Jan 2022 and Dec 2022 103
#19 (FOB gold* or FOBgold* or SENTiFIT):ti,ab,kw OR (JACK-arc* or JACKarc* or HM-JACK* or HM JACK* or HMJACK*):ti,ab,kw OR (OC Sensor* or OC-Sensor* or OCSensor* or Ceres or OC Pledia* or OC-Pledia* or OCPledia or OC-iO):ti,ab,kw OR (POC FIT QRG or POCFITQRG or immundiagnostik or IDK or turbifit or turbitube or quikread):ti,ab,kw OR (NS-Prime or NSPrime or NS-Plus):ti,ab,kw (Word variations have been searched)175
#20 #18 or #19 with Cochrane Library publication date Between Jan 2016 and Dec 2022 224

INAHTA (searched 13/12/2022)
Single word strings:
Faecal / fecal / colorectal / colon / cecal = 0 results

NIHR HTA programme website (searched 13/12/2022)
Searched website - only found a few blogs including references to the RECEDE study

PROSPERO searched 13/12/2022
This website only allows for simple searches:
Colorectal AND faecal (records added to PROSPERO since 1/1/22) = 15 results
Colorectal AND fecal = 20 results (including the 15 above)
fecal immunochemical test = 6 results
faecal immunochemical test = 6 results
faecal occult blood = 7 results
fecal occult blood = 11 results
FOBT = 8 results
MeSH Colorectal Neoplasms/ = 41 results
Faecal and test* = 51 results
Fecal and test* and cancer = 27 results

ClinicalTrials.gov searched 13/12/2022
(CTgov automatically expands the search to include synonyms and alternate spellings)

Colorectal cancer AND faecal = 341 results since 1/1/2016
Colorectal cancer AND FIT = 159 results since 1/1/2016
Colon cancer AND faecal = 90 results since 1/1/2016
Colon cancer AND FIT = 32 results “”
Rectal cancer AND FIT = 11 results “”
Rectal cancer AND faecal = 92 results “”

EU Trials Register (searched 13/12/2022)

0 results

WHO ICTRP (13/12/2022)

colon cancer OR colorectal cancer OR rectal cancer OR cecal cancer
AND
faecal OR fecal OR FIT or FOBT or iFOBT
32 results

B. Economic modelling search strategies

Cost-effectiveness and quality of life studies of FIT in patients with symptoms suggestive of colorectal cancer

Ovid MEDLINE(R) and Epub Ahead of Print, In-Process, In-Data-Review & Other Non-Indexed Citations, Daily and Versions <1946 to February 22, 2023>

1 f?ecal immunochemical test.mp. 1292
2 f?ecal occult blood.mp. 4479
3 f?ecal h?emoglobin.mp. 272
4 ((immunochromatographic or immuno-chromatographic or immunochem$ or immuno-chem$ or immunohistochem$ or immuno-histochem$ or immunol$ or immunoassay or immuno* assay or immunoturbidimetric or immunosorbent or elisa) adj4 (f?ecal or f?eces or stool or stools or FIT)).mp. 3658
5 (iFOBT or qFIT).mp. 214
6 or/1-5 7377
7 F?ecal h?emoglobin.ti,ab,ot,hw. 258
8 H?emoccult.ti,ab,ot,hw. 728
9 FOBT.ti,ab,ot,hw. 1440
10 7 or 8 or 9 2348
11 (f?ecal or f?eces or stool or stools).ti,ab,ot,hw. 214450
12 occult blood/ or occult blood.ti,ab,ot,hw. 8990
13 (test$ or measur$ or screen$ or exam$).ti,ab,ot,hw. 10821494
14 11 and 12 and 13 6077
15 6 or 10 or 14 8824
16 exp colorectal neoplasms/ 233170
17 exp cecal neoplasms/ 6078
18 ((colorect$ or rectal$ or rectum$ or colon$ or sigma$ or sigmo$ or rectosigm$ or bowel$ or anal or anus) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot,hw. 327674
19 CRC.ti,ab,ot. 44474
20 ((cecum or cecal or caecum or caecal or il?eoc?ecal or il?eoc?ecum) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 2778
21 (large intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 1841
22 (lower intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 34
23 16 or 17 or 18 or 19 or 20 or 21 or 22 339635
24 15 and 23 6007
25 (FOB gold$ or FOBgold$ or SENTiFIT).mp. 39
26 (JACK-arc$ or JACKarc$ or HM-JACK$ or HM JACK$ or HMJACK$).mp. 23

27 (OC Sensor$ or OC-Sensor$ or OCSensor$ or Ceres).mp. 373
28 (OC Pledia$ or OC-Pledia$ or OCPledia or OC-iO).mp. 0
29 (NS-Prime or NSPrime or NS-Plus).mp. 37
30 (POC FIT QRG or POCFITQRG).mp. 0
31 (immundiagnostik or IDK or turbifit or turbitube).mp. 126
32 quikread.mp. 20
33 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 6459
34 Economics/ 27492
35 exp "costs and cost analysis"/ 262760
36 Economics, Dental/ 1920
37 exp economics, hospital/ 25681
38 Economics, Medical/ 9240
39 Economics, Nursing/ 4013
40 Economics, Pharmaceutical/ 3095
41 (economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab. 1010903
42 (expenditure$ not energy).ti,ab. 36017
43 value for money.ti,ab. 2078
44 budget$.ti,ab. 34677
45 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 1174425
46 ((energy or oxygen) adj cost).ti,ab. 4690
47 (metabolic adj cost).ti,ab. 1676
48 ((energy or oxygen) adj expenditure).ti,ab. 28548
49 46 or 47 or 48 33866
50 45 not 49 1166602
51 letter.pt. 1208094
52 editorial.pt. 636950
53 historical article.pt. 369079
54 or/51-53 2193150
55 50 not 54 1126959
56 exp animals/ not humans/ 5094439
57 55 not 56 1053612
58 bmj.jn. 87108
59 "cochrane database of systematic reviews".jn. 16141
60 health technology assessment winchester england.jn. 1496
61 or/58-60 104745
62 57 not 61 1046870
63 33 and 62 1042
64 limit 63 to yr="2016 -Current" 388
65 quality-adjusted life years/ or quality of life/ 272427
66 (sf36 or sf 36 or sf-36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).ti,ab,ot. 29855
67 (sf6 or sf 6 or sf-6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).ti,ab,ot. 2562
68 (sf12 or sf 12 or sf-12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,ot. 7345
69 (sf6D or sf 6D or sf-6D or short form 6D or shortform 6D or sf six D or sfsixD or shortform six D or short form six D).ti,ab,ot. 980
70 (sf20 or sf 20 or sf-20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,ot. 448
71 (sf8 or sf 8 or sf-8 or short form 8 or shortform 8 or sf eight or sfeight or shortform eight or short form eight).ti,ab,ot. 728
72 "health related quality of life".ti,ab,ot. 55787
73 (Quality adjusted life or Quality-adjusted-life).ti,ab,ot. 16616

74 "assessment of quality of life".ti,ab,ot. 2154
75 (euroqol or euro qol or eq5d or eq 5d).ti,ab,ot. 15788
76 (hql or hrql or hqol or h qol or hrqol or hr qol).ti,ab,ot. 26757
77 (hye or hyes).ti,ab,ot. 75
78 health$ year$ equivalent$.ti,ab,ot. 40
79 (hui or hui1 or hui2 or hui3 or hui4 or hui-4 or hui-1 or hui-2 or hui-3).ti,ab,ot. 1904
80 (quality time or qwb or quality of well being or "quality of wellbeing" or "index of wellbeing" or "index of well being").ti,ab,ot,hw. 1127
81 (Disability adjusted life or Disability-adjusted life or health adjusted life or health-adjusted life or "years of healthy life" or healthy years equivalent or "years of potential life lost" or "years of health life lost").ti,ab,ot. 5764
82 (QALY$ or DALY$ or HALY$ or YHL or HYES or YPLL or YHLL or qald$ or qale$ or qtime$ or AQoL$).ti,ab,ot. 19557
83 (timetradeoff or time tradeoff or time trade-off or time trade off or TTO or Standard gamble$ or "willingness to pay").ti,ab,ot. 10702
84 15d.ti,ab,ot. 1923
85 (HSUV$ or health state$ value$ or health state$ preference$ or HSPV$).ti,ab,ot. 493
86 (utilit$ adj3 ("quality of life" or valu$ or scor$ or measur$ or health or life or estimat$ or elicit$ or disease$)).ti,ab,ot. 15634
87 (utilities or disutili$).ti,ab,ot. 9413
88 (Functional Assessment of Cancer Therapy$ or FACT-G).ti,ab,ot. 3041
89 (QLQ-C30 or QLQ-C-30 or EORTC QLQ$ or "European Organization for Research and Treatment of Cancer Quality of Life Questionnaire" or "EORTC Quality of Life Questionnaire").ti,ab,ot. 6633
90 or/65-89 342017
91 33 and 90 132
92 91 not 64 68

Embase <1974 to 2023 Week 07>

1 f?ecal immunochemical test.mp. 2444
2 f?ecal occult blood.mp. 7022
3 f?ecal h?emoglobin.mp. 470
4 ((immunochromatographic or immuno-chromatographic or immunochem$ or immuno-chem$ or immunohistochem$ or immuno-histochem$ or immunol$ or immunoassay or immuno* assay or immunoturbidimetric or immunosorbent or elisa) adj4 (f?ecal or f?eces or stool or stools or FIT)).mp. 6393
5 (iFOBT or qFIT).mp. 411
6 or/1-5 12159
7 F?ecal h?emoglobin.ti,ab,ot,hw. 454
8 H?emoccult.ti,ab,ot,hw. 989
9 FOBT.ti,ab,ot,hw. 2830
10 7 or 8 or 9 4155
11 (f?ecal or f?eces or stool or stools).ti,ab,ot,hw. 287846
12 occult blood/ or occult blood.ti,ab,ot,hw. 18594
13 (test$ or measur$ or screen$ or exam$).ti,ab,ot,hw. 14258724
14 11 and 12 and 13 10644
15 6 or 10 or 14 15206
16 exp colorectal cancer/ or colon cancer/ or rectum cancer/ 377137
17 exp cecum tumor/ 7397
18 ((colorect$ or rectal$ or rectum$ or colon$ or sigma$ or sigmo$ or rectosigm$ or bowel$ or anal or anus) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot,hw. 521146

19 CRC.ti,ab,ot. 73761
20 ((cecum or cecal or caecum or caecal or il?eoc?ecal or il?eoc?ecum) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 3624
21 (large intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 1821
22 (lower intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).ti,ab,ot. 44
23 16 or 17 or 18 or 19 or 20 or 21 or 22 537004
24 15 and 23 10244
25 (FOB gold$ or FOBgold$ or SENTiFIT).mp. 115
26 (JACK-arc$ or JACKarc$ or HM-JACK$ or HM JACK$ or HMJACK$).mp. 78
27 (OC Sensor$ or OC-Sensor$ or OCSensor$ or Ceres).mp. 798
28 (OC Pledia$ or OC-Pledia$ or OCPledia or OC-iO).mp. 0
29 (NS-Prime or NSPrime or NS-Plus).mp. 79
30 (POC FIT QRG or POCFITQRG).mp. 0
31 (immundiagnostik or IDK or turbifit or turbitube).mp. 431
32 quikread.mp. 56
33 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 11170
34 health-economics/ 35319
35 exp economic-evaluation/ 349167
36 exp health-care-cost/ 332717
37 exp pharmacoeconomics/ 227487
38 34 or 35 or 36 or 37 739026
39 (econom$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab. 1357302
40 (expenditure$ not energy).ti,ab. 49546
41 (value adj2 money).ti,ab. 2941
42 budget$.ti,ab. 46255
43 39 or 40 or 41 or 42 1401000
44 38 or 43 1760393
45 letter.pt. 1291591
46 editorial.pt. 764786
47 note.pt. 926965
48 45 or 46 or 47 2983342
49 44 not 48 1627860
50 (metabolic adj cost).ti,ab. 1840
51 ((energy or oxygen) adj cost).ti,ab. 5006
52 ((energy or oxygen) adj expenditure).ti,ab. 36873
53 50 or 51 or 52 42520
54 49 not 53 1619172
55 exp animal/ 30254802
56 exp animal-experiment/ 3050020
57 nonhuman/ 7373530
58 (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep).ti,ab,sh. 6395872
59 55 or 56 or 57 or 58 32471017
60 exp human/ 25120688
61 exp human-experiment/ 638565
62 60 or 61 25123083
63 59 not (59 and 62) 7349104
64 54 not 63 1465032
65 33 and 64 1817
66 limit 65 to yr="2016 -Current" 704
67 quality adjusted life year/ or quality of life index/ 37657

68 Short Form 12/ or Short Form 20/ or Short Form 36/ or Short Form 8/ 47910
69 "International Classification of Functioning, Disability and Health"/ or "ferrans and powers quality of life index"/ or "gastrointestinal quality of life index"/ 4281
70 (sf36 or sf 36 or sf-36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirtysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).ti,ab,ot. 49160
71 (sf6 or sf 6 or sf-6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).ti,ab,ot. 2885
72 (sf12 or sf 12 or sf-12 or short form 12 or shortform 12 or sf twelve or sftwelve or shortform twelve or short form twelve).ti,ab,ot. 11999
73 (sf6D or sf 6D or sf-6D or short form 6D or shortform 6D or sf six D or sfsixD or shortform six D or short form six D).ti,ab,ot. 1827
74 (sf20 or sf 20 or sf-20 or short form 20 or shortform 20 or sf twenty or sftwenty or shortform twenty or short form twenty).ti,ab,ot. 512
75 (sf8 or sf 8 or sf-8 or short form 8 or shortform 8 or sf eight or sfeight or shortform eight or short form eight).ti,ab,ot. 1186
76 "health related quality of life".ti,ab,ot. 82918
77 (Quality adjusted life or Quality-adjusted-life).ti,ab,ot. 26027
78 "assessment of quality of life".ti,ab,ot. 3479
79 (euroqol or euro qol or eq5d or eq 5d).ti,ab,ot. 29306
80 (hql or hrql or hqol or h qol or hrqol or hr qol).ti,ab,ot. 44711
81 (hye or hyes).ti,ab,ot. 162
82 health$ year$ equivalent$.ti,ab,ot. 41
83 (hui or hui1 or hui2 or hui3 or hui4 or hui-4 or hui-1 or hui-2 or hui-3).ti,ab,ot. 3863
84 (quality time or qwb or "quality of well being" or "quality of wellbeing" or "index of wellbeing" or index of well being).ti,ab,ot,hw. 1521
85 (Disability adjusted life or Disability-adjusted life or health adjusted life or health-adjusted life or "years of healthy life" or healthy years equivalent or "years of potential life lost" or "years of health life lost").ti,ab,ot. 7121
86 (QALY$ or DALY$ or HALY$ or YHL or HYES or YPLL or YHLL or qald$ or qale$ or qtime$ or AQoL$).ti,ab,ot. 33625
87 (timetradeoff or time tradeoff or time trade-off or time trade off or TTO or Standard gamble$ or "willingness to pay").ti,ab,ot. 16459
88 15d.ti,ab,ot. 2918
89 (HSUV$ or health state$ value$ or health state$ preference$ or HSPV$).ti,ab,ot. 760
90 (utilit$ adj3 ("quality of life" or valu$ or scor$ or measur$ or health or life or estimat$ or elicit$ or disease$)).ti,ab,ot. 25437
91 (utilities or disutili$).ti,ab,ot. 15477
92 (Functional Assessment of Cancer Therapy$ or FACT-G).ti,ab,ot. 5504
93 (QLQ-C30 or QLQ-C-30 or EORTC QLQ$ or "European Organization for Research and Treatment of Cancer Quality of Life Questionnaire" or "EORTC Quality of Life Questionnaire").ti,ab,ot. 13823
94 or/67-93 259151
95 33 and 94 239
96 95 not 66 112

Econlit <1886 to February 09, 2023>

1 f?ecal immunochemical test.mp. 0
2 f?ecal occult blood.mp. 12
3 f?ecal h?emoglobin.mp. 0
4 ((immunochromatographic or immuno-chromatographic or immunochem$ or immuno-chem$ or immunohistochem$ or immuno-histochem$ or immunol$ or immunoassay

or immuno* assay or immunoturbidimetric or immunosorbent or elisa) adj4 (f?ecal or f?eces or stool or stools or FIT)).mp. 5
5 (iFOBT or qFIT).mp. 3
6 or/1-5 15
7 F?ecal h?emoglobin.mp. 0
8 H?emoccult.mp. 2
9 FOBT.mp. 7
10 7 or 8 or 9 9
11 (f?ecal or f?eces or stool or stools).mp. 104
12 occult blood.mp. 12
13 (test$ or measur$ or screen$ or exam$).mp. 516772
14 11 and 12 and 13 12
15 6 or 10 or 14 16
16 ((colorect$ or rectal$ or rectum$ or colon$ or sigma$ or sigmo$ or rectosigm$ or bowel$ or anal or anus) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).mp. 151
17 CRC.mp. 93
18 ((cecum or cecal or caecum or caecal or il?eoc?ecal or il?eoc?ecum) adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).mp. 0
19 (large intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).mp. 0
20 (lower intestin$ adj3 (cancer$ or neoplas$ or oncolog$ or malignan$ or tumo?r$ or carcinoma$ or adenocarcinoma$ or sarcoma$ or adenom$ or lesion$)).mp. 0
21 16 or 17 or 18 or 19 or 20 214
22 15 and 21 15
23 (FOB gold$ or FOBgold$ or SENTiFIT).mp. 0
24 (JACK-arc$ or JACKarc$ or HM-JACK$ or HM JACK$ or HMJACK$).mp. 0
25 (OC Sensor$ or OC-Sensor$ or OCSensor$ or Ceres).mp. 17
26 (OC Pledia$ or OC-Pledia$ or OCPledia or OC-iO).mp. 0
27 (NS-Prime or NSPrime or NS-Plus).mp. 0
28 (POC FIT QRG or POCFITQRG).mp. 0
29 (immundiagnostik or IDK or turbifit or turbitube).mp. 0
30 quikread.mp. 0
31 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 32
32 15 or 31 33

Cochrane search – already run as part of clinical SLR in Dec 2022
Re-ran to find new records added between Dec 2022 and Feb 2023:
Search Name:
Date Run: 23/02/2023 16:18:46
Comment:

ID Search Hits
#1 (fecal immunochemical test* or faecal immunochemical test*):ti,ab,kw (Word variations have been searched) 500
#2 (fecal occult blood or faecal occult blood):ti,ab,kw (Word variations have been searched) 1091
#3 (fecal hemoglobin or faecal hemoglobin or fecal haemoglobin or faecal haemoglobin):ti,ab,kw (Word variations have been searched) 299
#4 ((immunochromatographic or immuno-chromatographic or immunochem* or immuno-chem* or immunohistochem* or immuno-histochem* or immunol* or immunoassay or immuno* assay or immunoturbidimetric or immunosorbent or elisa) near/4 (fecal or faecal or feces or faeces or stool or stools or FIT)):ti,ab,kw (Word variations have been searched) 1047

#5 (iFOBT or qFIT):ti,ab,kw (Word variations have been searched) 37
#6 (Hemoccult or haemoccult):ti,ab,kw (Word variations have been searched) 129
#7 (FOBT):ti,ab,kw (Word variations have been searched) 412
#8 ((fecal or feces or faecal or faeces or stool or stools)):ti,ab,kw AND (occult blood):ti,ab,kw AND (test* or measur* or screen* or exam*):ti,ab,kw (Word variations have been searched) 1158
#9 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 2202
#10 MeSH descriptor: [Colorectal Neoplasms] explode all trees 10857
#11 MeSH descriptor: [Cecal Neoplasms] explode all trees 27
#12 ((colorect* or rectal* or rectum* or colon* or sigma* or sigmo* or rectosigm* or bowel* or anal or anus) near/3 (cancer* or neoplas* or oncolog* or malignan* or tumo* or carcinoma* or adenocarcinoma* or sarcoma* or adenom* or lesion*)):ti,ab,kw (Word variations have been searched) 26712
#13 (CRC):ti,ab,kw (Word variations have been searched) 5145
#14 ((cecum or cecal or caecum or caecal or ileocecal or ileocecum or ileocaecal or ileocaecum) near/3 (cancer* or neoplas* or oncolog* or malignan* or tumo* or carcinoma* or adenocarcinoma* or sarcoma* or adenom* or lesion*)):ti,ab,kw (Word variations have been searched) 247
#15 (large intestin* near/3 (cancer* or neoplas* or oncolog* or malignan* or tumo* or carcinoma* or adenocarcinoma* or sarcoma* or adenom* or lesion*)):ti,ab,kw (Word variations have been searched) 175
#16 (lower intestin* near/3 (cancer* or neoplas* or oncolog* or malignan* or tumo* or carcinoma* or adenocarcinoma* or sarcoma* or adenom* or lesion*)):ti,ab,kw (Word variations have been searched) 183
#17 #10 or #11 or #12 or #13 or #14 or #15 or #16 27565
#18 #9 and #17 with Cochrane Library publication date Between Jan 2022 and Dec 2022 106
#19 (FOB gold* or FOBgold* or SENTiFIT):ti,ab,kw OR (JACK-arc* or JACKarc* or HM-JACK* or HM JACK* or HMJACK*):ti,ab,kw OR (OC Sensor* or OC-Sensor* or OCSensor* or Ceres or OC Pledia* or OC-Pledia* or OCPledia or OC-iO):ti,ab,kw OR (POC FIT QRG or POCFITQRG or immundiagnostik or IDK or turbifit or turbitube or quikread):ti,ab,kw OR (NS-Prime or NSPrime or NS-Plus):ti,ab,kw (Word variations have been searched) 177
#20 #18 or #19 with Cochrane Library publication date Between Jan 2016 and Dec 2022 229
#21 #18 or #19 with Cochrane Library publication date Between Dec 2022 and Feb 2023 5

Cost Effectiveness Analysis (CEA) Registry (Internet)
https://research.tufts-nemc.org/cear4/Home.aspx
Date searched: 23.2.23
Records found: 195

Search term (basic search)	Records found
Colonoscopy	50
Computed tomographic colonography	2
CT colonography	2
Coloscopy	0
Sigmoidoscopy	4
Magnetic resonance imaging	64
MRI	60
CT scan	12
CAT scan	0
TOTAL	**195**

Appendix 2: Conversion of sensitivity and specificity to TP, TN, FP, FN

When the absolute number of diagnostic counts (TP, TN, FN, FP) were not reported by a study, but data for the total number of patients, the total number of positive cases, sensitivity and specificity were available, the count data were calculated using the equations below.

TP = sensitivity x number of positive cases

FN = (1 - sensitivity) x number of positive cases

FN = (1 - specificity) x (total number - number of positive cases)

TN = specificity x (total number - number of positive cases)

Appendix 3: QUADAS scoring scheme and scores with reasons for all studies

a) Scoring Scheme

Domain 1: Patient selection
Was a consecutive or random sample of patients enrolled? *o Score yes if states consecutive or random* *o Score no if states another method of patient sampling/selection* *o Score unclear if unclear*
Was a case-control design avoided? *o Score yes if not case control* *o Score no if case control* *o Score unclear if unclear* *(there should be no case control studies in the included studies, but please double check)*
Did the study avoid inappropriate exclusions? *Score yes if the study only excluded bypass symptom patients* *Score no if the study has made inappropriate exclusions e.g. on basis of having had a colonoscopy, taking certian medications, having blood disorders, or on the basis of eventually being diagnosed with IBD (list not exhaustive)* *Score unclear if it is unclear*
Risk of bias summary score: Could the selection of patients have introduced bias? Low/High/Unclear THIS IS A SUMMARY SCORE BASED ON ANSWERS TO QUESTIONS 1 TO 3. Score Low if all domains are Yes Score High if one or more domain is No Anything in between score Unclear
Applicability summary score: Is there concern that the included patients and settings do not match the review question? *Score low if the study selected all patients presenting to primary care with symptoms of CRC as listed in DG30 and NG12. If the study recruited a wider population, i.e. patients who do not meet these criteria, please state unclear risk (wider).* *Score High if the study missed some of the primary care patients, e.g. if only those referred to colonoscopy were recruited (unless all primary care NG12/DG30 are referred to secondary care)* Low/High/Unclear
Domain 2: Index test(s)
Were the index test results interpreted without knowledge of the results of the reference standard? *o Score yes if index test was interpreted blind to the reference standard or the index test*

was clearly interpreted before the reference standard was known, e.g. FIT before colonoscopy *o Score no if results of reference standard were already known e.g. FIT done after colonoscopy* *o Score unclear if unclear*
If a threshold was used, was it pre-specified? *o Score yes if pre-specified cut off values were used (validation study) e.g. one or a range of cut-offs reported such as 10, 20 , 50 , 100 ug/g and these were not chosen on the basis of having the highest accuarcy* *o Score no if cut-off values were fitted to the data (derivation study) e.g. cut-off with highest accuracy reported* *o Score unclear if unclear* *NB if study reports both the highest precision cut-off, and several other "round number" cut-offs, score yes/no*
Risk of bias summary score: Could the conduct or interpretation of the index test have introduced bias? Low/High/Unclear **THIS IS A SUMMARY SCORE BASED ON ANSWERS TO QUESTIONS 1 TO 2.** **Score Low if all domains are Yes** **Score High if one or more domain is No** **Anything in between score Unclear**
Applicability summary score: Is there concern that the index test, its conduct, or interpretation differ from the review question? *We may need to ask Sally to help us know what is normal practice, so for now we will just extract data* Low/High/Unclear
Domain 3: Reference Standard
Is the reference standard likely to correctly classify the target condition? ***Please note limitations/test type against score*** *Score **Yes** if all patients received either colonoscopy or CT colonography (CTC)* *Score **No** if the reference standard was not full colonic imaging (see yes criteria) for all patients* *Score **Unclear** if its unclear*
Were the reference standard results interpreted without knowledge of the results of the index test? *In the case of tiered testing, this is likely not to be the case.* o Score yes if the reference standard was interpreted blind to the index test or the reference standard was clearly interpreted before the index test was known. o Score no if the results of the index test were known, e.g. where patients were referred on the basis of a FIT result. o Score unclear if unclear
Risk of bias summary score: Could the reference standard, its conduct, or its interpretation have introduced bias?

Low/High/Unclear
(THIS IS A SUMMARY SCORE BASED ON ANSWERS TO QUESTIONS 1 TO 2)

Applicability summary score: Is there concern that the target condition as defined by the reference standard does not match the review question?
*Score **low risk** if the target condition is CRC*
*Score **high risk** if the target is not just CRC*
*Score **unclear** if the target condition is unclear*
NB: all studies should score low risk
NB: we are not scoring for AA and IBD
Low/High/Unclear

The reference standard may be free of bias, but the target condition that it defines may differ from the target condition specified in the review question.

Domain 4: Flow and timing

Was there an appropriate interval between index test(s) and reference standard?
Score low risk if all patients received colonoscopy and this was conducted within 3 months of the index test, or if some patients received records follow-up, this should be for a minimum of 3 months
Score high risk if colonoscopies were not conducted within 3 months, or follow-up is for less than 3 months but more than 12 months
Score unclear if the time intervals were unclear
NB: likely most won't report time interval for colonoscopy

Yes/No/Unclear

Did all patients receive a reference standard?
*Score **yes** if all patients got a reference standard, even if these were different (see next question)*
*Score **no** if a partial verification reference standard: only some participants get any reference standard, e.g. those who test negative at FIT don't get followed up or any further tests (these studies should be excluded)*
*Score **unclear** if it is unclear who received the reference standard*

Did patients receive the same reference standard?

*The following score "**no**":*
Complete index test-dependent differential verification reference standard: participants get a different reference standard according to the index test result, e.g. FIT positive get colonoscopy, FIT negative get records follow-up
Differential verification dependent on other known or unknown factors: participants get a different reference standard according to some known or unknown factors, e.g. those with clinical signs or symptoms proceed to colonoscopy regardless of FIT, whilst reminder get records follow-up

*The following score "**yes**":*
All received the same reference standard, e.g. all get colonoscopy

Were all patients included in the analysis? *Score **yes** if all patients who were recruited/enrolled into the study were included in the analysis or if an acceptable explanation (i.e missing at random) is provided for any discrepancy* *Score **no** if there are participants excluded from the analysis and no/concerning explanation is given for any discrepancy* *Score **unclear** if insufficient information is given to assess whether any patients were excluded from the analysis.*
Risk of bias summary score: Could the patient flow have introduced bias? Low/High/Unclear (THIS IS A SUMMARY SCORE BASED ON ANSWERS TO QUESTIONS 1 TO 4) Score Low if all domains are Yes Score High if one or more domain is No Anything in between score Unclear

b) Reasons for all scores

Table 74: HM-JACKarc studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores

	1. Was a consecutive or random sample of patients enrolled?	2.Was a case-control design avoided?	3. Did the study avoid inappropriate exclusions?	RoB: Could the selection of patients have introduced bias?	Applicability: Is there concern that the included patients and settings do not match the review question?	1. Were the index test results interpreted without knowledge of the	2. If a threshold was used, was it pre-specified?	RoB: Could the conduct or interpretation of the index test have introduced bias?	Applicability: Is there concern that the index test, its conduct, or interpretation differ from the review question?	1. Is the reference standard likely to correctly classify the target condition?	2. Were the reference standard results interpreted without knowledge of the results of the index test?	RoB: Could the reference standard, its conduct, or its interpretation have introduced bias?	Applicability: Is there concern that the target condition as defined by the reference standard does not match the review question?	1. Was there an appropriate interval between index test(s) and reference standard?	2. Did all patients receive a reference standard?	3. Did patients receive the same reference standard?	4. Were all patients included in the analysis?	RoB: Could the patient flow have introduced bias?
Benton 2022[46]	no (participation voluntary)	yes	no (subgroup, participation	high	high	yes	yes	low	Low	yes	unclear	unclear	low	unclear	yes	yes	yes (some missing data,	low

			voluntary)														explanations given, no reason to assume not random)	
Chapman 2021[48]	yes	yes	no	high	High	yes	yes	low	low	yes (follow-up, implied all colonoscopy or imaging)	unclear	unclear	low	unclear	yes	unclear - because some may have received more than just colonoscopy	yes (missing data lack of follow-up)	unclear
Cunin 2020[74]	yes	yes	yes	low	high	yes	yes	low	high	no	unclear	high	low	unclear	yes	no	yes	high
D'Souza	yes	yes	no	high	Unclear - some	yes	yes	low	low	yes	unclear	unclear	low	unclear	yes	yes	yes	low

2020a[49]					may be missed													
D'Souza 2021a[75], D'Souza 2021c[17]	unclear	yes	yes	unclear	high	yes	yes	low	low	yes	unclear	unclear	low	unclear	yes	yes	yes	unclear
D'Souza 2021b[76]	unclear	yes	no (colonoscopy)	high	high	yes	yes	low	low	yes	unclear	unclear	low	unclear	yes	yes	yes	unclear
Elbeltagi 2022[50]	Yes - states "all"	Yes	Unclear	Unclear	High - patients on 2WW who had COL/CTC so mix of pts	Yes	Yes	Low	Low	Yes (colonoscopy/CTC)	Unclear	Unclear	Low	Unclear	Yes	No, unclear how determined	No only those who had col/CTC	High

Farrugia 2020[77]	unclear	yes	no	high	high	yes	yes	low	low	yes	unclear	unclear	low	unclear	yes	yes	yes	unclear
Faux 2022[51]	Unclear	Yes	No, rectal bleeding and anaemia excluded	High	High	Yes	Yes	Low	Low	No	Unclear	High	Low	Unclear	Yes	No, dependent on FIT and presenting symptom	Yes	High
Gerrard 2023[78]	Yes, states consecutive	Yes	No, IDA on it's own was not a referral criterion; only those with colonoscopy/CTC	High	High	Yes	Yes	Low	Low	No	Unclear, but secondary care investigation received not based on FIT	High	Low	Yes - states median time to diagnosis	Yes	yes all either col or CTC	No - some excluded as triaged away from imaging	High
Godber 2016[53]	yes	yes	no	high	high	yes	yes	low	low	yes	unclear	unclear	low	unclear	yes	yes	yes	low

Johnstone 2022a[55]	yes (database)	yes	yes	low	low	yes	yes	low	low	no	unclear	unclear	low	unclear	yes (some imaging, some follow-up)	no (some imaging, some follow-up)	yes	high
Johnstone 2022b[56]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Mac Donald 2022[59]	Yes - states consecutive	Yes	Yes	Low	Unclear risk (wider and narrower)	yes	Yes	Low	Low	No - not all had colonoscopy/CTC, unclear what 1113 had	Unclear - FIT used to define test in secondary care but unclear if interpreted blind	High	Low	High - records follow-up for >2 years could allow CRC to emerge; unclear how long between FIT and imaging/ secondary care	yes	No - Differential verification dependent on other known or unknown factors, as 1113	Yes	High

														appointment		unknown		
Mowat 2021[64] & 2019[63]	Unclear - uptake of FIT by GPs was slow, unclear who was missed in early adoption period	Yes	Yes (CRC results); No (IBD/HRA results)	Unclear	Unclear (it's not entirely clear that ALL NG12 and DG30 patients were included, but it's likely the case. Though use of FIT was rolled	Yes	Yes	Low	Low	No - some did not get full imaging	No	High	Low	Unclear - time interval for colonoscopy not stated	Yes	No - Differential verification dependent on other known or unknown factors: some patients with FIT<10 referred for COL, some with FIT>10 not referred.	Yes	High

					out slowly, so not clear if this was due to more GPs adopting, or GPs becoming more confident to use in all symptoms. If the latter, may skew distribution towar													

					ds more or less "serious" symptoms)													
Nicholson (2018)[66]	Yes, states consecutive	Yes	Yes	Low	High, only DG30	Yes	Yes	Low	High - some patients had two test results, and if either was positive the test was jud	No - records follow up and imaging	Unclear - not stated	High	Low - states adenocarcinoma, but no other types of CRC were reported in the "significant bowel pathology" category	No - follow-up was for 21 and 23 months which may have allowed for the emergency of CRC since the index test	Yes	No - unclear if completely dependent on index test	Unclear - not clear if more patients were offered FIT but did not complete it	High

									ged as positive									
Nicholson (2020)[29a]	Low	Yes	Yes	Low	High - only DG30	Yes	Yes	Low	High - some patients had two test results, and if either was positive the test was judged as pos	No - records follow up and imaging	Unclear - not stated	High	Low	Yes - minimum of 6 months, analyses show that longer follow up did not significantly alter the sens/spec	Yes	No - unclear if completely dependent on index test	Unclear - some excluded due to not long enough follow-up but thisis unlikely to introduce bias, but not clear if patients missing for	High

									itive								other reasons also	
Tang 2022[69]	yes	yes	no (needed both FIT and colonoscopy)	high	high	yes	yes	low	low	yes	unclear	unclear	low	Unclear	yes	yes	yes	Unclear
Turvill 2021[85]	No - convenience sample	Yes	Yes - all patients recruited accounted for and reasons for exclusions not likely to skew selection	High	High - some patients outside of NG12 criteria	Yes	Yes /No - some are derived optimal cut off, some are round numbers	High /Low	Low	No - some did not get full imaging	Yes - states blind	High	Low	Unclear - not reported	Yes	No - Differential verification dependent on other known or unknown factors	Yes - some missing but unlikely to skew results	High
Turvill 2018[84]	no (consent to study,	yes	no (colonoscopy or imaging)	high	high	yes	no	high	high	no	yes	high	low	unclear	yes	no	yes (reasons for exclus	high

	convenience series)															ion given, no reason to assume not random)	
Withrow 2022[7] [0a]	yes - states all	yes	no - only patients with five most common core blood tests available were included	High	High	yes	yes	low	low	no - follow up	unclear	High	Low	Unclear	yes (follow-up)	no (some no imaging)	no - unclear why some missing	High

[a] Nicholson 2020 and Withrow 2022 include some of the same patients, but it was not clear if the same methodology was used in both studies to select patients and conduct follow-up, so scores are provided for each study based on the information given for that study

Table 75: OC-Sensor studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores

	1. Was a consecutive or random sample of patients enrolled?	2.Was a case-control design avoided?	3. Did the study avoid inappropriate exclusions?	RoB: Could the selection of patients have introduced bias?	Applicability: Is there concern that the included patients and settings do not match the review question?	1. Were the index test results interpreted without knowledge of the results of the reference standard?	2. If a threshold was used, was it pre-specified?	RoB: Could the conduct or interpretation of the index test have introduced bias?	Applicability: Is there concern that the index test, its conduct, or interpretation differ from the review question?	1. Is the reference standard likely to correctly classify the target condition?	2. Were the reference standard results interpreted without knowledge of the results of the index test?	RoB: Could the reference standard, its conduct, or its interpretation have introduced bias?	Applicability: Is there concern that the target condition as defined by the reference standard does not match the review question?	1. Was there an appropriate interval between index test(s) and reference standard?	2. Did all patients receive a reference standard?	3. Did patients receive the same reference standard?	4. Were all patients included in the analysis?	RoB: Could the patient flow have introduced bias?
Archer 2022	yes (all referred)	yes	no - excluded those without colonos	high	high - those ref to 2WW, +	yes	yes	low	low	yes	unclear	unclear	low	unclear	yes	yes	no	unclear

			copy or FIT test		colonoscopy + Fit test													
Ayling 2019	unclear	yes	no, only colonoscpy/CT included	high	high	yes	yes	low	low	yes (assume CT scan is CTC)	yes ("faecal sample for immunological measurement of haemoglobin and their blood count parameters were analysed using an artificial intelligence (AI) flaggin	low	low	unclear	yes	no	no	high

											g system; these results were not made available for patient management")							
Ball 2022 (personal communicaton)	Yes	Yes	Yes	High	High	Yes	Yes	Low	Low	No - records follow up	Unclear	High	High	No - unclear if colonscopy done within 3 months; records follow up was at least 18 month	Yes	No - unclear if complete or partial differential bais	Yes	High

														s, which could allow for CRC to emerge after the index test				
Ball 2022	Yes	Yes	Yes	Low	Low	Yes	Yes	Low	Low	No - records follow up	Unclear	High	High	No - unclear if colonscopy done within 3 months; records follow up was at least 18 months, which	Yes	No - unclear if complete or partial differential bais	Yes	High

														could allow for CRC to emerge after the index test				
Bujanda 2018	yes	yes	no (all colonoscopy; occasional aspirin users)	high	high	yes	yes	low	low	yes	yes	low	low	unclear	yes	yes	yes	unclear
Cama 2022	Yes - states consecutive	Yes	No - excluded IDA and rectal bleeding	High	High	Yes	Yes	Low	Low	No	Unclear	High	Low	Yes - 12 months follow up	Yes	No - some will have had imaging, some will have had records	Yes	High

Chapman 2021	yes	yes	no	high	High	yes	yes	low	low	yes (follow-up, implied all colonoscopy or imaging)	unclear	unclear	low	unclear	yes	unclear - because some may have received more than just colonoscopy	yes (missing data lack of follow-up)	unclear
Georgiou Delisle 2022	Yes, states all	yes	No, some patients sent to alternative pathways (rectal bleeding or CLASP) which may not be available in all areas of	High	High	Yes	Yes	Low	Low	No - follow-up for some patients	Unclear	Unclear	Low	Yes - no patients excluded on basis of colonoscopy not having been performed at 6 month	Yes	Yes - Complete index test-dependent differential verification reference standard: participants get a differe	Yes	Low

			England											s follow-up .		nt reference standard according to the index test result		
Hunt 2022	Yes - states "all"	Yes	No - only some areas tested those with rectal bleeding	High	High - rectal bleeding not representative; due to covid there is an unrepresentative mix of patients in the total cohort	Yes	Yes	Low	Low	Unclear - not clear how many patients had colonoscopy/CTC	Unclear	Unclear	Low	Unclear - doesn't state time period between FIT and imaging; doesn't state how long recrods follow up for	Yes	Unclear - doesn't state how many had col/CTC or not	Yes	Unclear

Juul 2018	Yes - states "all"	Yes	No, excluded some patients that meet NG12 high risk, but not all	High	High	Yes	Yes	Low	Low	No - not all had colonoscopy/CTC	No - states doctors performing colonoscopy not blind	High	low	Unclear	yes	no (some did not get diagnostic investigation)	yes	High
Laszlo 2021	Unclear - does not state "all" or "consecutive"	Yes	no - excluded those without a definitive diagnosis	High	High	Yes, states blind	yes	low	low	Yes	yes - states blind	Low	Low	Unclear, does not state interval	Yes	No - Differential verification dependent on other known or unknown factors	No - excluded patients whithout definitive diagnosis, which may have introduced bias if these patients are system	High

																	atically different	
Maclean 2021a	Yes - states all	yes	Yes	low	High - 2WW	Yes - supplement	yes	low	low	No - not all col	No - used to triage patients to reference standard	High	Low	Unclear	Unclear (not clear if those ref back to GP were followed up)	No	No - some excluded due to frailty, cancellation, etc	High
Morales-Arraez 2018	yes	yes	no (needed both FIT and colonoscopy)	high	high	yes	yes	low	low	yes	unclear	unclear	low	Unclear	yes	yes	yes	Unclear
Mowat 2016	Yes	Yes	No - only patients who had a	High	High	Yes	Yes/ No - LoD selected	High (LoD); Lo	Low	Yes	Yes - All clinicians and endosc	Low	Low	Yes - Patients triaged to	Yes	Yes	No - pts w/o colonoscopy exclud	High

			colonoscopy were included, based on secondary care triage				as >/= 10 had FNs - therefore LoD high risk, 10 low risk	w (10 ug/g)			opists were blind to the faecal test results.			endoscopy were investigated within 6 weeks.			ed & 5 pts missing from analysis without explanation	
Pin Vieto	Unclear - just states inclusion criteria, but not that were consecutive or "all"	Yes	Yes - excluded pts with CRC in prior 2 years	Unclear	Unclear - not enough information provided to score	Yes	Yes	Low	Low	No - some did not get full imaging	No - states performed as "part of their medical treatment"	High	Low	Unclear - not clear how quickly colonoscopy was done; Follow-up of records for 2 years	yes	No - Complete index test-dependent differential verification reference standard: participants	Unclear	High

																get a different reference standard according to the index test result, e.g. FIT positive get colonoscopy, FIT negative get records follow-up (2 years)		
Rodriguez-Alonso 2018,	Unclear	Yes	No, excluded premen	High	High - only secondary care	Unclear	Yes	Unclear	Low (NBOC-Sen	Yes (colonoscopy)	Unclear	Unclear	Low	Unclear - does not	yes	yes	Unclear who was exclud	Unclear

Rodriguez-Alonso 2019			opausal women, some others		referrals, some referred from secondary care not primary care (included due to reporting anaemia data)				sor MICRO)					state interval			ed based on incomplete tests etc	
Crooks 2023, Bailey J 2021a	yes (all logged)	yes	no	high	high	yes	yes	low	low	no	unclear	high	low	unclear	yes	no	no	high

Table 76: FOB-Gold studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores

	1. Was a consecutive or random sample of patients enrolled?	2.Was a case-control design avoided?	3. Did the study avoid inappropriate exclusions?	RoB: Could the selection of patients have introduced bias?	Applicability: Is there concern that the included patients and settings do not match the review question?	1. Were the index test results interpreted without knowledge of the results of the reference standard?	2. If a threshold was used, was it pre-specified?	RoB: Could the conduct or interpretation of the index test have introduced bias?	Applicability: Is there concern that the index test, its conduct, or interpretation differ from the review question?	1. Is the reference standard likely to correctly classify the target [illegible]	2. Were the reference standard results interpreted without knowledge of the results of the index test?	RoB: Could the reference standard, its conduct, or its interpretation have introduced bias?	Applicability: Is there concern that the target condition as defined by the reference standard does not match the review question?	1. Was there an appropriate interval between index test(s) and reference standard?	2. Did all patients receive a reference standard?	3. Did patients receive the same reference standard?	4. Were all patients included in the analysis?	RoB: Could the patient flow have introduced bias?
Benton 2022[46]	no (participation voluntary)	yes	no (subgroup, participation	high	high	yes	yes	low	Low	yes	unclear	unclear	low	unclear	yes	yes	yes (some missing data, explanations	low

			voluntary)														given, no reason to assume not random)	
MacLean 2022a[62]	Unclear - does not state consecutive or all	Yes	No - only included those with definitive diagnosis, which may exclude a spectrum of patients	High	High	Yes, states blind	Yes	Low	Low	Yes	Yes, states blind	Low	Low	Unclear, does not state interval	yes	No - Differential verification dependent on other known or unknown factors	No - excluded patients without definitive diagnosis, non-return of sample, pack not received, DNAs all excluded	High
Schwettmann 2022[68]	Unclear	Yes	Unclear	Unclear	High	Yes	Yes	Low	Low - but note different analyser	Yes	Unclear	Unclear	Low	Yes - collected 3-7 days before colonoscopy	yes	Yes	Unclear	Unclear

Table 77: QuikRead go studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores

	1. Was a consecutive or random sample of patients enrolled?	2. Was a case-control design avoided?	3. Did the study avoid inappropriate exclusions?	RoB: Could the selection of patients have introduced bias?	Applicability: Is there concern that the included patients and settings do not match the review question?	1. Were the index test results interpreted without knowledge of the results of the reference standard?	2. If a threshold was used, was it pre-specified?	RoB: Could the conduct or interpretation of the index test have introduced bias?	Applicability: Is there concern that the index test, its conduct, or interpretation differ from the review question?	1. Is the reference standard likely to correctly classify the target	2. Were the reference standard results interpreted without knowledge of the results of the index test?	RoB: Could the reference standard, its conduct, or its interpretation have introduced bias?	Applicability: Is there concern that the target condition as defined by the reference standard does not match the review question?	1. Was there an appropriate interval between index test(s) and reference standard?	2. Did all patients receive a reference standard?	3. Did patients receive the same reference standard?	4. Were all patients included in the analysis?	RoB: Could the patient flow have introduced bias?
Maclean 2021b[61]	Unclear - does not state	Yes	No - only included those	High	High	Yes, states	Yes	Low	Low	Yes	Yes, states	Low	Low	Unclear, does not state	Yes	No - Differential verificat	No - excluded patients without .	High

	consecutive or all		with definitive diagnosis, which may exclude a spectrum of patients			blind					blind			interval		ion dependent on other known or unknown factors	definitive diagnosis, non-return of sample, pack not received, DNAs all excluded	
Tsapournas 2020[83]	No - convenience sample	Yes	No, only included those with colonoscopy	High	High - includes referrals from primary and secondary care	Yes - states blind	Yes	Low	Low	Yes	Yes - states blind	Low	Low	Yes, <30 days (states tests >30 days excluded)	Yes	Yes	No, excluded patients who did not have a colonoscopy but doesn't detail why	High

Table 78: NS-Prime study: ScHARR's assessment of risk of bias and applicability, with reasons for scores

	1. Was a consecutive or random sample of patients enrolled?	2.Was a case-control design avoided?	3. Did the study avoid inappropriate exclusions?	RoB: Could the selection of patients have introduced bias?	Applicability: Is there concern that the included patients and settings do not match the review question?	1. Were the index test results interpreted without knowledge of the	2. If a threshold was used, was it pre-specified?	RoB: Could the conduct or interpretation of the index test have introduced bias?	Applicability: Is there concern that the index test, its conduct, or interpretation differ from the review question?	1. Is the reference standard likely to correctly classify the target	2. Were the reference standard results interpreted without knowledge of the results of the index test?	RoB: Could the reference standard, its conduct, or its interpretation have introduced bias?	Applicability: Is there concern that the target condition as defined by the reference standard does not match the review question?	1. Was there an appropriate interval between index test(s) and reference standard?	2. Did all patients receive a reference standard?	3. Did patients receive the same reference standard?	4. Were all patients included in the analysis? .	RoB: Could the patient flow have introduced bias?
Benton 2022[4][6]	no (participation voluntary)	yes	no (subgroup, participation voluntary)	high	high	yes	yes	low	Low	yes	unclear	unclear	low	unclear	yes	yes	yes (some missing data, explanations given,	low

																	no reason to assume not random)	

Table 79: FOB-Gold studies: ScHARR's assessment of risk of bias and applicability, with reasons for scores

	1. Was a consecutive or random sample of patients enrolled?	2.Was a case-control design avoided?	3. Did the study avoid inappropriate exclusions?	RoB: Could the selection of patients have introduced bias?	Applicability: Is there concern that the included patients and settings do not match the review question?	1. Were the index test results interpreted without knowledge of the results of the reference standard?	2. If a threshold was used, was it pre-specified?	RoB: Could the conduct or interpretation of the index test have introduced bias?	Applicability: Is there concern that the index test, its conduct, or interpretation differ from the review question?	1. Is the reference standard likely to correctly classify the target condition?	2. Were the reference standard results interpreted without knowledge of the results of the index test?	RoB: Could the reference standard, its conduct, or its interpretation have introduced bias?	Applicability: Is there concern that the target condition as defined by the reference standard does not match the review question?	1. Was there an appropriate interval between index test(s) and reference standard?	2. Did all patients receive a reference standard?	3. Did patients receive the same reference standard?	4. Were all patients included in the analysis?	RoB: Could the patient flow have introduced bias?
Sieg 1999	Unclear	Yes	No, those going to colonoscopy	High	High	Yes - states	Hb: yes Hb/Hp: no	Hb: Low Hb/Hp: High	Yes - no data relating to the equivalen	Yes	Yes - states	Low	Low	Unclear	Yes	Yes	Unclear (N recruited not .	Unclear

						blind			ce of this test and the commercial version has been presented in symptomatic patients		blind						reported)	

Appendix 4: Statistical methods for the evidence synthesis

The statistical model is briefly described following the notation in Jones et al. [38] True disease status is assumed to be known through application of a perfect gold standard test. Populations without and with CRC are indexed by $j = 1, 2$ respectively. Each study, *i*, reports estimates of sensitivity and specificity, or directly reports count data, at T_i distinct thresholds. Test results above a given threshold are considered positive.

The observed count data is modelled using multinomial likelihoods, reparametrized as conditional binomial distributions for computational convenience. The model assumes that some transformation, $g()$ of the continuous test results in population j of study i has a logistic distribution with mean μ_{ij} and scale parameter σ_{ij}. In our analyses we pre-specify $g(\ \ ) = \log_e()$. Jones et al. describe the more flexible (but computationally intensive) case where $g(\ \ )$ is in the set of Box-Cox transformation, defined by a parameter which is estimated alongside other model parameters.

Within study model

The probability of a positive test result at threshold C_{it} in population j of study i is

$$logit(pr_{ijt}) = \frac{(\mu_{ij} - g(C_{it}))}{\sigma_{ij}} \quad (1)$$

For $j = 1$ we have pr_{i1t} (the false positive rate, FPR=1-sensitivity) and for $j = 2$ we have pr_{i2t} (the true positive rate, TPR=sensitivity)

Between study model

The study specific location (μ_{ij}) and scale (σ_{ij}) parameters are modelled as random effects. Across studies, μ_{ij}, has mean $m_{\mu j}$ and standard deviation $\tau_{\mu j}$.while log (σ_{ij}) has mean $m_{\sigma j}$ and standard deviation $\tau_{\sigma j}$

Different options for the correlation structure between these four sets of random effects are described in Jones *et al.*: i) Full correlation matrix, ii) structured correlations matrix iii) independence model with the four sets of random effects assumed to be independent of each other. Models with a structured correlation matrix and assuming independence were explored. Including additional parameters for between-study correlations did not improve the model fit according to the DIC (see Appendix 5, Table 81), therefore the simpler independence model was used for all main analyses.

Prior distributions are required for the four hyperparameters : $m_{\mu j}$, $\tau_{\mu j}$, $m_{\sigma j}$ $\tau_{\sigma j}$ for $j = 1,2$. For analyses with sufficient sample data, standard reference priors as used in Jones *et al.*were used. Normal (0, 10^2) prior distribution were given to all means ($m_{\mu j}, m_{\sigma j}$), and Uniform (0,5) prior for between study standard deviations ($\tau_{\mu j}, \tau_{\sigma j}$)

For analyses with small numbers of contributing studies, informative priors were used for the between study standard deviations. These were informed by fitting log-normal distributions were fitted to posterior samples from the analyses of all test types together. This was considered to be a conservative option. A truncation was also applied, based on the 95th centile of the posterior distribution. Parameter values for all analyses are provided in Table 80

Table 80: Parameters used to inform priors for syntheses with less than 5 studies.

Parameter	CRC outcomes (S=28)			AA outcomes (S=9)			IBD outcomes (S=9)		
	mean	sd	truncation	mean	sd	truncation	mean	sd	truncatio
$\tau_{\mu 1}$	0.2698	0.1543	1.703	0.2859	0.3432	2.478	0.4532	0.3561	3.0
$\tau_{\mu 2}$	-0.8489	0.2995	0.67	-0.4517	0.8583	2.174	0.3827	0.8695	3.9
$\tau_{\sigma 1}$	-0.8863	0.1559	0.538	-0.5228	0.415	1.247	-0.4273	0.4077	1.34
$\tau_{\sigma 2}$	-1.4215	0.2693	0.368	0.152	0.5711	3.079	-0.8792	0.9243	1.33

Appendix 5: Additional meta-analysis results and NPV and PPV results

a) **Additional meta-analysis results:** Model fit for all analyses shown in Table 81. Differences in DIC across the two correlations structures (structured and independent) were minimal. Since including additional parameters for between-study correlations did not improve the model fit according to the DIC, the simpler model structure was preferred and all analyses presented in the report use the model of Jones *et al.*[38] with the four sets of random effects assumed to be independent of each other (independence model).

Table 81: Meta-analysis model fit statistics

Tests in analysis	Populations	Studies	Correlation structure	Model fit		
				$\overline{D}$	pD	DIC
CRC outcomes						
All	All	28	S	6718.40	72.91	6791.31
	All	28	I	6711.25	71.78	6783.03
	1,2,3	13	I	1780.45	31.63	1812.08
	1	8	I	1541.81	21.73	1563.54
	2	5	I	173.87	NaN	173.87
	2	5	I	175.77	10.26	186.02
	3	3	I	107.51	NaN	107.51
HM JACKArc	All	16	S	5297.88	41.45	5339.32
	All	16	I	5296.11	40.46	5336.56
	1,2,3	9	I	648.54	19.29	667.83
	1	5	I	495.25	12.69	507.93
	2	4	I	143.50	8.11	151.60
	3	2	I	39.00	Nan	39.00
OC Sensor	All	11	S	1395.86	28.05	1423.92
	All	11	I	1394.99	28.71	1423.70
	1,2,3	4	I	1112.37	11.18	1123.55
	1	3	I	1044.86	9.35	1054.21
FOB Gold	All	3	I	114.56	NaN	Nan
AA outcomes						
All tests	All	9	I	308.81	23.41	332.21
HM-JACKarc	All	6	I	240.87	13.86	254.74
OC Sensor	All	2	I	38.67	6.02	44.69
IBD outcomes						
All tests	All	9	I	286.38	23.97	310.36
HM-JACKarc	All	6	I	220.37	14.64	235.01
OC Sensor	All	2	I	38.27	6.12	44.39

DUAL FIT						
All tests	All	4	I	63.87	8.03	71.90

S: Structured correlation matrix. I: Independence model

b) PPV and NPV for selected thresholds, from the synthesised sensitivity and specificity, or individual studies where no synthesis was performed

Table 82: PPV and NPV for selected thresholds, from the synthesised sensitivity and specificity, or individual studies where no synthesis was performed

The PPV and NPV have been calculated at selected, available thresholds. The prevalence used was as used in the model for the whole cohort, i.e. based on the meta-analysed values from the clinical review for studies of type 1.

Test	Prevalence	Threshold (µg/g)	Sensitivity	Specificity	PPV	NPV
CRC Outcomes						
HM-JACKarc	0.028	LoD (2)	0.959	0.651	0.07	1.00
	0.028	LoQ (7)	0.914	0.796	0.11	1.00
	0.028	10	0.895	0.828	0.13	1.00
	0.028	20	0.847	0.879	0.17	1.00
	0.028	50	0.758	0.926	0.23	0.99
	0.028	100	0.67	0.949	0.27	0.99
OC-Sensor	0.028	Lowest reported value (4)	0.942	0.627	0.07	1.00
	0.028	7	0.918	0.723	0.09	1.00
	0.028	10	0.898	0.776	0.10	1.00
	0.028	20	0.847	0.856	0.14	0.99
	0.028	50	0.75	0.925	0.22	0.99
	0.028	100	0.653	0.959	0.31	0.99
FOB Gold	0.028	Lowest reported value (2)	0.969	0.652	0.07	1.00
	0.028	7	0.93	0.775	0.11	1.00
	0.028	10	0.912	0.803	0.12	1.00
	0.028	20	0.864	0.851	0.14	1.00
	0.028	50	0.769	0.899	0.18	0.99
	0.028	100	0.67	0.926	0.21	0.99
QuikRead go	0.028	Lowest reported value (10)	0.929	0.701	0.08	1.00
	0.028	20	NR	NR	NR	NR
	0.028	50	NR	NR	NR	NR
	0.028	100	0.714	0.946	0.28	0.99
NS-Prime	0.028	Lowest reported value (3)	0.857	0.319	0.03	0.99
	0.028	7	NR	NR	NR	NR

	0.028	10	0.714	0.836	0.11	0.99
	0.028	20	NR	NR	NR	NR
	0.028	50	NR	NR	NR	NR
	0.028	100	0.571	0.973	0.38	0.99
IDK Hb	0.028	Lowest reported value (2)	0.87	0.881	0.17	1.00
IDK Hb/Hp	0.028	Lowest reported value (2)	0.826	0.808	0.11	0.99
AA Outcomes						
HM-JACKarc	0.02	LoD (2)	0.591	0.559	0.03	0.98
	0.02	LoQ (7)	0.547	0.685	0.04	0.98
	0.02	10	0.532	0.719	0.04	0.98
	0.02	20	0.509	0.779	0.05	0.99
	0.02	50	0.498	0.848	0.07	0.99
	0.02	100	0.487	0.882	0.09	0.99
OC-Sensor	0.02	Lowest reported value (4)	0.939	0.468	0.04	1.00
	0.02	7	0.846	0.707	0.06	0.99
	0.02	10	0.732	0.822	0.09	0.99
IBD Outcomes						
HM-JACKarc	0.027	LoD (2)	0.868	0.57	0.05	0.99
	0.027	LoQ (7)	0.801	0.699	0.07	0.99
	0.027	10	0.776	0.733	0.07	0.99
	0.027	20	0.723	0.792	0.09	0.99
	0.027	50	0.723	0.792	0.09	0.99
	0.027	100	0.592	0.891	0.13	0.99
OC-Sensor						
	0.027	Lowest reported value (4)	0.67	0.464	0.03	0.98
	0.027	7	0.598	0.703	0.05	0.98
	0.027	10	0.551	0.819	0.08	0.99

Appendix 6: Worked example relating to potential bias of excluding FIT negative patients

Worked example of the impact of enriching a study with FIT positives, and of using an imperfect reference standard

1. Data only for those going for colonoscopy (NG12 +DG30 FIT+ve)

Assume we have a study which recruited patients who had been referred to secondary care under recommendations from NG12 and DG30. Some were referred on the basis of NG12 symptoms alone without a FIT, and some were referred on the basis of having DG30 symptoms, and a positive FIT test. The study reports the following test characteristics

	FIT+	FIT-	**Total**

CRC+	100	10	110
CRC-	1000	4000	5000
Total	1100	4010	5110

Sensitivity: 100/110 = 90.9 %
Specificity = 4000/5000 = 80.0%

2. **Data for all patients presenting in primary care (i.e., including DG30 FIT-ves)**

If we assume that for every FIT+ in primary care that gets referred to secondary care, there were 7 FIT negative that did not get referred, and if we assume that approximately 10% of the patients in the study were referred on the basis of a FIT positive test in primary care, then:
Number of patients referred on the basis of a FIT positive in primary care = 5110*0.1 = 511
Number of FIT tests not referred from primary care because they were negative = 511*7 = 3,577
If we then also assume that amongst the negatives, the PPV was 0.24% (this is taken from Ball *et al.*),[45] then 9 CRC were missed (3577*0.24%).

Therefore the number of TPs and FPs stay the same, but:
TNs have an additional 3577-9 = 3,568 patients, which should be added to the TNs from the original sample: 3568+4000 = 7568
FNs have an additional 9, which should be added to the FNs from the original sample: 10+9 = 19

	FIT+	FIT-	**Total**
CRC+	100	19	**119**
CRC-	1000	7568	**8568**
Total	**1100**	**7587**	**8687**

Sensitivity: 100/119 = 84.0%
Specificity = 7568/8568 = 88.3%

3. **What is the effect of underestimating the true number of FNs by including DG30 FIT -ve with records follow-up?**

We think that follow up probably underestimates the true number of FNs and therefore overestimates the number of TNs, because it is an imperfect reference standard. So if we assume that there are 1, 10, or 20 FNs missed (that is 5%, 34% and 51% of CRCs missed), what happens to sensitivity and specificity?

This table adds in the FNs and subtracts them from the TNs (since with an imperfect reference standard, the missed FNs would have been counted as TNs)

	FIT+	FIT- (baseline)	FIT- (1 extra FN)	FIT- (10 extra FNs)	FIT- (20 extra FNs)

CRC+	100	19	20	29	39
CRC-	1000	7568	7567	7558	7548
	1100	**7587**	**7587**	**7587**	**7587**

The sensitivity and specificities are therefore:

Sensitivity 0 FN missed	0.840336	84.0%
Specificity 0 FN missed	0.883287	88.3%
Sensitivity 1 FN missed	0.833333	83.3%
Specificity 1 FN missed	0.883273	88.3%
Sensitivity 10 FN missed	0.775194	77.5%
Specificity 10 FN missed	0.88315	88.3%
Sensitivity 20 FN missed	0.719424	71.9%
Specificity 20 FN missed	0.883014	88.3%

The following table compares Type 4 studies and the "true numbers" (as calculated under an assumption that the number of FIT negatives remaining in primary care is 1 FIT+: 7 FIT-, and for various assumptions about the reference standard).

Scenario		**Sens**	**Spec**	**Interpretation (compare scenario sens/spec to "true numbers" estimate)**
Type 4 studies	Ref to Col pts only	90.9 %	80.0%	Type 4 studies overestimate sensitivity and underestimate specificity
Type 1 studies, not accounting for imperfect reference standard of follow-up	All presenting in primary care, 7 FIT negatives excluded for each FIT positive, PPV of negative FIT 0.24%	84.0%	88.3%	
"True numbers", under various assumptions about reference standard	All presenting to primary care, 7 FIT negatives excluded for each FIT positive, PPV of negative FIT 0.24% **Assume 1 FN missed by follow-up**	90.6	85.1	Type 4 studies underestimate specificity Type 1 studies also underestimate specificity
	Assuming 10 FN missed by follow-up	88.2	85.1	Type 4 overestimate sensitivity and underestimate specificity Type 1 studies overestimate sensitivity, little effect on specificity
	Assuming 20 FN missed by follow-up	85.6	85.1	Type 4 overestimate sensitivity and underestimate specificity

				Type 1 studies overestimate sensitivity, little effect on specificity

Conclusion

In our example, excluding DG30 FIT negatives (some type 4 studies) will **overestimate sensitivity** and **underestimate specificity** compared to **all other estimates, even where the reference standard misses some cases of CRC.**

We only looked at a limited set of assumptions, and the extent of the issue may vary according to these assumptions, and according to threshold used to define a positive test.

Appendix 7: Clinical review: Table of excluded studies with rationale

Studies excluded from the review of diagnostic test accuracy, with reasons

Reason	Number of studies excluded	References
Analytical performance	9	88, 127-134
Crossover, no new data or superseded	14	90, 135-147
Editorial, comment, letter	7	148-154
Incorrect population	40	155-194
Insufficient data to calculate DTA /data ambiguous/not DTA study	10	195-204
Outcome Not CRC or CRC only	4	205-208
Not English language	1	209
Not FIT or in-scope test	6	86, 210-214
Ongoing study or systematic review	10	215-224
Systematic review or review	13	25, 26, 92, 225-234
Threshold not reported	2	235, 236

Appendix 8: Diagnostic test accuracy data entering the statistical synthesis

This appendix will be provided as an addendum following submission of the report.

Appendix 9: Test uptake and repeat tests data from secondary care

Author, year	Analyser	FIT provided in	N with CRC/ N analysed (%)	Invalid/ test failure rates	Test uptake /non-return	Repeat tests
D'Souza 2020a[49]	HM JACKarc analytical system	Secondary care	12/298 (4.03%)		FIT not returned: 416/800 (52%)	
Gerrard 2023[78]	HM-JACKarc	Secondary care	135/3426 (3.05%)	Clinician considered FIT inappropriate, or emergency presentation predated FIT postage: 207/3074 (6.7%)	FIT not returned: 493/3074	
D'Souza 2021a[75]	HM JACKarc analytical system	Secondary care	421/9822 (4.29%)	FIT sample inadequate n=183/13219 (1.4%) FIT test >14 days old n=147/13219 (1.1%)	7907/21126 (37.4%) did not return FIT	
Faux 2022[51]	HM-JACKarc	Some primary, some secondary	6/175 (3.43%)		FIT not returned: 17/175 (8.9%)	
Godber 2016[53]	HM JACKarc analyser	Secondary care	11/484 (2.27%)		FIT not returned: 402/909 (44.2%)	
Tang 2022[69]	HM-JACKarc system	Secondary care	20/603 (3.32%)		FIT not returned: 13/280[a] (4.8%)	
Turvill 2021[85]	HM JACKarc	Secondary care	151/5040 (3.00%)	FIT sample incorrect: 49	did not return FIT: 1564/6864 = 22.8% Returned FIT after investigations: 60/6864 = 0.9%	
Laszlo 2021[58]	OC-Sensor iO	Secondary care	90/3596 (2.50%)	Not viable n=129/4676 (2.8%) Samples not received after patient returned it n=261/4676 (5.6%)		

Maclean 2021a[60]	OC-Sensor PLEDIA	Secondary care	12/358 (3.35%)	hospitalised n=5/391 (1.3%) cancelled n=5/391 (1.3%) FIT not received n=5/391 (1.3%)	FIT not: n=18/391 (4.6%)	
Multiple tests						
Benton 2022[46]	OC Sensor PLEDIA	Secondary care	7/233 (3.00%)	30/291 (10.3%)		
Chapman 2021[48]	HM JACKarc + HM JACKarc analyser OC-Sensor DIANA	Secondary care	38/732 (5.19%)	overall return rate for at least 1 device was 82.6%		
MacLean 2022a[62] Maclean 2021b[61]	QuikRead Go FOB Gold			QuikRead go: No error readings	Pack not received n=5 did not provide sample n=178	

[a]Patients who had FIT as first modality

Appendix 10: Review of economic evaluations and HRQoL studies: table of excluded studies with rationale and quality assessment of studies included

Review of economic evaluations studies

Author	Year	Reason for exclusion	Comment
Arasaradnam *et al*[122].	2020	Editorial paper or comments	-
Berger *et al.*[237]	2016	Not in the right population (asymptomatic patients)	-
Farkas *et al.*[238]	2023	Different type of study (not economic evaluation)	Also only available as an abstract
Gendia *et al.*[239]	2021	No full text available	Abstract only of model based on NICEFIT study
Hijos *et al.*[240]	2018	Does not include FIT	Also only available as an abstract, but the full text was also retrieved (Lue *et al.* 2020, presented below)
Seehra *et al.*	2021	No full text available	Study in subgroup of symptomatic >60yo with change in bowel habit (CiBH). Full text seems to be Khasawneh *et al.* 2022 (presented below).
Kearsey *et al.* [241]	2021	No full text available	-
Kearsey *et al.* [242]	2021	Not full economic evaluation (cost study/analysis)	Analysis based on retrospective clinical study in UK, but only evaluates costs, not the impact on LYs or HRQoL. Outcome is cost per cancer detected, but no model.
Khasawneh *et al.*[243]	2022	Not full economic evaluation (cost study/analysis)	Analysis based on retrospective clinical study in UKwith patients >=60yo with CiBH comparing patients referred to CTC before and after the introduction of FIT. However, it evaluates only costs, but not the impact on LYs or HRQoL. Economic impact analysis calculated by multiplying the number of CTCs saved by its unit cost.
Law *et al.*[244]	2016	Not in the right population (patients already diagnosed with colon polyps)	-
Lobo *et al.*[245]	2020	Not in the right population (patients already diagnosed with Crohn's disease)	Also, does not include FIT
Lue *et al.*[246]	2020	Does not include FIT	Analysis based on prospective clinical study in Spain (location outside scope of this review). It does not include FIT, population is of patients referred to secondary care (not primary), and cost-effectiveness outcome is cost per correctly diagnosed patient.
Padula *et al.*[247]	2017	Does not include FIT	Model based on patient clincial case (used decision tree to model the consequences of a patient receiving COL or not (not on primary care patients, wrong population).
Petersen *et al.*[248]	2020	Different type of study (not economic evaluation)	
Rodriguez-Alonso *et al.*[249]	2020	Not full economic evaluation (cost study/analysis)	Evaluated different FIT strategies (FIT-Gastrocopy, simple FIT and sequential FIT). based on prospective study in Spain of patients with iron deficiancy anemia only.

Sadeghi *et al.*	2016	Not in the right population (asymptomatic patients)	Also only abstract available, but the abstract indicates to be a study in screening patients (wrong population)
Saing *et al.*[250]	2018	Not in the right population (patients with confirmed carcinoma)	Study to evaluate ontrast-enhancedmagnetic resonance imaging in already diagnosed CRC patients
Smith *et al.*[251]	2019	Not in the right population (asymptomatic patients)	Study in screening patients (wrong population). The study looks at 1 FIT sample vs 2 samples.
Yu *et al.*[252]	2021	Not in the right population (asymptomatic patients)	SLRclinical review of screening models
Brar *et al.*[253]	2022	Not full economic evaluation (cost study/analysis)	Also only abstract available, and does not include FIT (mentions the total costs of testing on its abstract)
Braun *et al.*[254]	2020	Different type of study (not economic evaluation)	Analysed data from Swiss Health Interview Survey 2012 for associations between FOBT and colonoscopy testing and health insurance type
Coury *et al.*[255]	2022	Not full economic evaluation (cost study/analysis)	It looks at the costs of FIT and FOBT tests using a survey with primare care practices in US
Fisher *et al.*[256]	2022	Not full economic evaluation (cost study/analysis)	Estimates the costs of events in low GI screening and post-screening patients in US and does not include FIT.
Fisher *et al.*[257]	2022	Not full economic evaluation (cost study/analysis)	Also, study in screening patients (wrong population).
Paszat *et al.*[258]	2021	Not full economic evaluation (cost study/analysis)	Estimates the overall health care costs of 1 year following diagnosis after screening from Canada population-wide administrative database.
Pelitari *et al.*[259]	2021	Different type of study (not economic evaluation)	Study in screening population that evaluates the uptake of FIT compared with FOBT kits and its effects (uses FIT threshold of 120 mg/g). It includes costs estimates.
Subramanian *et al.*[260]	2019	Not full economic evaluation (cost study/analysis)	Study in screening population. Evaluates costs of US screening programs, the effect of large versus small programs on clinical and nonclinical costs, controlling for factors such as geographic location and type of screening test used.
Veettil et al[261]	2021	Not in the right population (secondary prevention in patients already diagnosed with adenomas)	-

Review of HRQoL studies

Author	Year	Reason for exclusion	Comment
Bobridge *et al.*[262]	2014	Does not report EQ-5D estimates	Study didn't collect EQ-5D and in screening patients. But it did include FIT.
Bobridge *et al.*[263]	2012	Only abstract available	Study didn't collect EQ-5D
Ferrari *et al.*[264]	2016	Only abstract available	Study didn't collect EQ-5D
Kapidzic *et al.*[265]	2012	Wrong population (not symptomatic CRC patients in primary care)	In screening Dutch patients. FIT value of ⩾50 ng ml−1 was considered positive. Provide EQ-5D score for FIT+ vs FIT-.
Kirkoen *et al.*[266]	2016	Does not report EQ-5D estimates	Study only collected SF-12 and Hospital Anxiety and Depression scale (HADS), alsopopulation is screening patients from South-East Norway
Kirkoen *et al.*[267]	2016	Does not report EQ-5D estimates	Study only collected SF-12 and HADS, also population is screening patients from South-East Norway
Kirkoen B*et al.*[268]	2014	Only abstract available	Also study only collected collected SF-12 and HADS in screening patients, comparing the psychological effects to receiving an invitation for colorectal cancer screening with either FS or FIT
Mahabaleshwarkar *et al.*[269]	2013	Study not available	No access to the full text
Marshall *et al.*[270]	2009	Does not report EQ-5D estimates	Also study done with physicians and general population, and related to screening
Miles *et al.*[271]	2015	Does not report EQ-5D estimates	Collected only FACT-C and intervention was FOBT in screening patients
Mountifield [272]*et al.*	2013	Only abstract available	Study collected only SF-36, Multidimensional Locus of Control, and the Speilberger State and Trait Scale. Population of patients with IBD for whom colonoscopy is indicated for the detection of Colorectal Cancer (CRC)
Van Dam *et al.*[273]	2010	Only abstract available	Results of the EQ-5D in terms of point estimates not provided in the abstract (not able to check which analyses were performed) and study in screening patients
Vermeer *et al.*[274]	2020	Does not report EQ-5D estimates	Does not present EQ-5D and it is in screening, but presents SF-36 for patients with and without cancer before and after colonoscopy
Wattchow *et al.*[275]	2006	Does not report EQ-5D estimates	Study collected only SF-12 and HADS
Whynes *et al.* [276]	1994	Does not report EQ-5D estimates	Study from 1994 and collected only the Health Measurement Questionnaire (HMQ) and the Nottingham Health Profile
Yamada *et al.*[277]	2017	Does not report EQ-5D estimates	Study collected SF-8 (inclding a visual analogue scale (VAS)) in 100 consecutive Japanese screening patients with positive FOBT who underwent screening colonoscopy. Looked into colonoscopy-related pain in QOL.
Medina-Lara *et al.*. [93]	2020	A full economic evaluation study	It was included in the review of economic evaluations
Westwood *et al.*.[92]	2017	A full economic evaluation study	It was included in the review of economic evaluations

Quality assessment of studies included in the review of economic studies – Drummond checklist[91]

	Westwood et al (2017)	Medina-Lara et al (2020)
Study design		
1. Was the research question stated?	Yes	Yes
2. Was the economic importance of the research question stated?	Yes	Yes
3. Was/were the viewpoint(s) of the analysis clearly stated and justified?	Partially, mentioned but justification not provided	Partially, mentioned but justification not provided
4. Was a rationale reported for the choice of the alternative programmes or interventions compared?	Yes	Yes
5. Were the alternatives being compared clearly described?	Yes	Partially, description of how the interventions include the risk tools is unclear
6. Was the form of economic evaluation stated?	Yes	Yes
7. Was the choice of form of economic evaluation justified in relation to the questions addressed?	Yes	Yes
Data collection		
8. Was/were the source(s) of effectiveness estimates used stated?	Yes	Yes
9. Were details of the design and results of the effectiveness study given (if based on a single study)?	N/A (accuracy based on systematic review)	N/A (sources from multiple studies)
10. Were details of the methods of synthesis or meta-analysis of estimates given (if based on an overview of a number of effectiveness studies)?	Yes	Unclear
11. Was/were the primary outcome measure(s) for the economic evaluation clearly stated?	Yes	Yes
12. Were the methods used to value health states and other benefits stated? Yes	Yes	Yes
13. Were the details of the subjects, from whom valuations were obtained, given?	N/A; utilities from literature	Yes
14. Were productivity changes (if included) reported separately? NA	N/A	N/A
15. Was the relevance of productivity changes to the study question discussed?	N/A	N/A
16. Were quantities of resources reported separately from their unit cost?	Yes	Yes
17. Were the methods for the estimation of quantities and unit costs described?	Yes	Yes
18. Were currency and price data recorded?	Yes	Yes
19. Were details of price adjustments for inflation or currency conversion given?	Yes	unclear
20. Were details of any model used given?	N/A	Yes
21. Was there a justification for the choice of model used and the key parameters on which it was based?	Yes	Yes
Analysis and interpretation of results		
22. Was the time horizon of cost and benefits stated?	lifetime mentioned	Yes

	exact number of years unclear	
23. Was the discount rate stated?	Yes	Yes
24. Was the choice of rate justified?	No	No
25. Was an explanation given if cost or benefits were not discounted?	N/A	N/A
26. Was/were the details of statistical test(s) and CIs given for stochastic data?	Distribution chosen and parameters reported (including reference)	Yes
27. Was the approach to sensitivity analysis described?	Yes	Yes
28. Was the choice of variables for sensitivity analysis justified?	Yes	Yes
29. Were the ranges over which the parameters were varied stated?	Yes	Partially (some not stated but reference provided)
30. Were relevant alternatives compared? (That is, were appropriate comparisons made when conducting the incremental analysis?)	Yes	Yes
31. Was an incremental analysis reported?	Yes	Yes
32. Were major outcomes presented in a disaggregated form as well as aggregated form?	Yes	Yes
33. Was the answer to the study question given?	Yes	Yes
34. Did conclusions follow from the data reported?	Yes	Yes
35. Were conclusions accompanied by the appropriate caveats?	Yes	Yes
36. Were generalisability issues addressed?	Unclear	Yes

Appendix 11: EAG survey collected from EAG's clinical advisors

The following questionnaire was circulated via email to EAG's clinical advisors.

Questionnaire:

DAP50 Questions to clinical advisors

Thank you for agreeing to participate in DAPSO, and for all your help so far. As part of our assessment, we will be building an economic model to assess the most clinically and cost-effective way to use FIT as an adjunct to clinical assessment to guide referral of a symptomatic population to the suspected CRC pathway. To help us build and populate the model, we have some questions that we hope you can help us with, based on your clinical experience and any data you may be aware of.

* indicates required question

1. What is your name? *

Data on the population

2. i. Are you aware of any empirical data sources which describe the population (age distribution, deprivation, sex, etc) presenting to primary care with symptoms suggestive of CRC?

3. ii. Are you aware of any data describing the population (age distribution, deprivation, sex, etc) currently referred for 2WW?

4. iii. Are you aware of any data describing the population (age distribution, deprivation, sex, etc) currently referred for 18WW?

5. iv. Are you aware of any data describing the prevalence of disease (CRC, IBD, advanced adenomas [AA]) in these populations?

Current practice and capacity issues

Part of the assessment's aim is to reduce the number of referrals to the 2WW pathway who do not have significant bowel pathology.

6. i. What guidelines are you currently following - NG12/DG30, BSG or other?

7. ii. What are the approximate current waiting times for 2WW referrals and for 18WW referrals?

8. iii. Are waiting times improving or worsening or staying the same? Why do you think this is happening?

9. iv. What reduction (either number, **N,** or percentage, p%) in lower GI referrals would be required to enable target wait times to be met? Is this reduction in all referrals or just urgent 2WW referrals? If the reduction was **N,** would a reduction of N/2 put the delay midway between the current and the target wait time?

10. v. What would be the advantages of moving some referrals to 18WW from 2WW? E.g., Does demand on one pathway impact on waiting times for the other?

11. vi. What is considered significant bowel pathology (AA, non-advanced adenomas, IBD, diverticulitis... etc)? Which conditions would be detected by each of the diagnostic modalities used to investigate CRC as per Table in question 3.ii?

Urgent 2WW referral under current care

12. i. Currently, we believe patients may receive COL, CTC, CT or capsule COL in their urgent referral appointment. Is this correct? Are there any other investigations they may receive? E.g., may patients also receive an appointment with a specialist without imaging and be discharged to primary care, or referred to another pathway?

13. ii. How is the investigation received at 2WW referral currently determined? Is there significant heterogeneity? What proportion of patients receive each investigation? Please provide estimates for COL, CTC, CT, capsule COL and any other options (e.g. discharge to primary care; referral to another pathway) for age groups <60, 60-69, 70-79 and 80+ based on your clinical expertise, or if available indicate sources of real world

data. Do you anticipate this may change in the near future? Please explain why you anticipate the change, e.g. investigation likely to be phased out and/or capacity increased.

NB you won't be able to fill in the table due to the limitations of google forms. Suggested answer format for age groups <60; 60-69; 70-79; and 80+ is:

COL: 90%; 85%;
60%; 0% CTC:
CT:
Capsule

	Proportion of patients receiving each investigation			
Investigation	Age <60	Age 60-69	Age 70-79	Age 80+
COL				
CTC				
CT				
Capsule COL				
Other (please define e.g. discharge to primary care; referral to another pathway)				

CT:
other:

14. iii. We have heard that FIT is used to decide who should receive CTC versus COL where there is a shortage of COL due to capacity. Please describe how this was/is done, if known.

15. iv. Under DG30/NG12 where FIT is not completed, or under other pathways where FIT is not completed, how was/is the choice of investigation determined?

16. v. If COL capacity was not an issue then would FIT still be useful to inform secondary care decisions regarding whether COL or CTC is most appropriate?

17. vi. Would patients with bypass symptoms (defined in the NICE scope for DAP50 as palpable rectal or anal mass, or anal ulceration) still receive FIT alongside their referral? Would FIT be given in primary care or secondary care? How would the FIT results be used?

Non-urgent (18WW?) referral under current care

18. i. Is the non-urgent pathway an 18WW pathway?

---

19. ii. Is this pathway currently for those not suspected of having cancer? If so, what conditions are generally suspected? Under what circumstances/signs/symptoms are patients referred to this pathway?

---

20. iii. In your practice, is this pathway further subdivided into "routine" and "urgent"? If so, what determines this subdivision? Or do you triage patients on this pathway in some other way?

---

21. iv. How is the investigation received at (routine and urgent)18WW referral determined?

---

22. v. What investigations might patients receive at 18WW referral? Please provide estimates for COL, CTC, GT, capsule COL and any other options (e.g. discharge to primary care; referral to another pathway) for age groups <60, 60-69, 70-79 and 80+ based on your clinical expertise, or if available indicate sources of real world data.

NB you won't be able to fill in the table due to the limitations of google forms. Suggested answer format for age groups <60; 60-69; 70-79; and 80+ is:

COL: 90%; 85%;
60%; 0% CTC:
GT:
Capsule
GT:
other:

	Proportion of patients receiving each investigation			
Investigation	Age <60	Age 60-69	Age 70-79	Age 80+
COL				
CTC				
CT				
Capsule COL				
Other (please define)				
Add more rows as needed...				

Management pathways we will be modelling

We will be modelling the following ways to use FIT to guide referral of a symptomatic population to the suspected CRC pathway:

- **NG12/DG30** diagnostic pathways
- Guidance issued by NHS England advising use of the **ACPGBI/BSG** guidelines
- **Intervention 1:** A FIT strategy which uses a single FIT threshold to determine management pathways. A range of different single FIT thresholds will be considered (e.g., 2, 10, 20, 50 and 100 µg/g of faeces).
- **Intervention 2:** A FIT strategy using two thresholds which we are calling t1 (upper threshold) and t2 (lower threshold) to determine management pathways. A range of different pairs of FIT thresholds will be considered (e.g., 10 and 100µg/g). See Figure 1.

The next questions are designed to help us understand what safety netting and the intermediate risk pathways might look like.

Pathways for the "intermediate" risk group (t1-t2)

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

23. i. What management options could be appropriate for an intermediate group with FIT score between thresholds t1 and t2? E.g. non-urgent referral to secondary care (and what would it include: appointment with specialist, imaging, etc.); other options? In a situation where non-urgent referrals were seen within 18 weeks would this be an appropriate management strategy to consider for the intermediate group?

    Would management differ according to the threshold selected? If so, in what way? for example, what would be appropriate for a FIT score between 10 and 100, vs a FIT score between 10 and 50?

Safety Netting pathways

In the assessment we are defining the term 'safety netting' to refer to the patient pathways for persons whose primary care appointment does not result in direct referral to the 2WW pathway. For intervention 1 those with FIT score <t1 receive safety netting, for intervention 2 those with FIT score <t2 receive safety netting, and for NG12/DG30 patients not referred receive safety netting. Persons not completing FIT may also receive safety netting. Note that safety netting may include strategies for diagnosing other gastrointestinal conditions, and further monitoring for colorectal or other types of cancer.

24. i. Do you have or are you aware of any empirical data relating to safety netting?

---

25. ii. We believe safety netting to be very heterogeneous, but largely comprising:
    Referral to 2WW
    Non-urgent referral
    (18WW) Watch and wait
    Referral to another diagnostic
    pathway A repeat FIT
    Use of eRS to guide next steps
    Use of automated text messages to a) encourage return of test in non-returners or b) representation if symptoms worsen or persist, or new symptoms emerge
    Use of software to list non-completers/patients for follow-up

Are these correct? Are there additional options we have not listed?

---

26. iii. eRS: we heard from some clinicians that this is a two-way communication system with secondary care to help guide referral in primary care, but from others that it isn't used. Please could you comment on how eRS is used in your practice, and across England if known? If eRS is not used for advice and guidance between primary and secondary care, please describe whether and how this is done in your practice.

---

27. iv. What currently happens to patients who do not return their FIT? Is there heterogeneity in this? What should happen to patients who do not return their FIT in your opinion?

---

28. v. What proportion of patients do you anticipate would receive each of the following safety netting pathways, assuming a **low intensity safety netting for patients with FIT<10:**

- Watch and wait
- Repeat **FIT**
-eRS
- Non-urgent referral
- Referral to 2WW
- Referral to another diagnostic pathway

Add more rows as needed…

---

29. vi. What proportion of patients do you anticipate would receive each of the following safety netting pathways, assuming a **high intensity safety netting, for patients with FIT<100:**

- Watch and wait
- Repeat **FIT**
-eRS
- Non-urgent referral
- Referral to 2WW
- Referral to another diagnostic pathway

Add more rows as needed…

30. vii. What proportion of patients do you anticipate would receive each of the following safety netting pathways, assuming **current care (for patients not receiving 2WW)**

- Watch and wait
- Repeat FIT
- eRS
- Non-urgent referral
- Referral to 2WW
- Referral to another diagnostic pathway
Add more rows as needed…

---

31. viii. What is an estimate of the overall cost of safety netting per person, under an assumption of each of low intensity, high intensity and current care?

Efficacy of Safety Netting

32. i. How much **diagnostic delay** would you expect for a patient (with a FIT result below the threshold) who receives management via safety netting? Please provide an estimate for each method below (input a range if you prefer), based on your clinical expertise (or if available indicate sources of real world data) for patients with an **underlying health state** of each of CRC, IBD and AA. For example, for a patient with underlying IBD who was not referred to secondary care on the 2WW pathway, how many weeks of diagnostic delay might a patient incur with safety netting: watch and wait?

NB you won't be able to fill in the table due to the limitations of google forms. Suggested response format:

Watch and wait: CRC estimate; IBD estimate; AA estimate
Repeat FIT: CRC estimate; IBD estimate; AA estimate
eRS: etc...
Referral to another diagnostic pathway: etc...
Referral to 2WW:

Add more rows as needed…

NB for patients referred to the **non-urgent referral pathway** (18WW), we suggest that we will assume the delay to be the waiting time for non-urgent referral - does this seem reasonable?

Safety Netting pathway followed	Estimated diagnostic delay (weeks)		
	Underlying health state: CRC	Underlying health state: IBD	Underlying health state: AA
Watch and wait			
Repeat FIT			
eRS			
Non-urgent referral	*	*	*
Referral to another diagnostic pathway			
Add pathways if required:			

*It is suggested that we will assume the delay to be the waiting time for non-urgent referral

33. ii. Would a delay in diagnosis of IBD be likely to result in a significant impact on disease progression (and resulting impact on Qol, disease severity, outcomes, treatment costs)?

34. iii. Would a delay in diagnosis of any other significant pathology (e.g. diverticular disease) be likely to result in a significant impact on disease progression (and resulting impact on Qol, disease severity, outcomes, treatment costs)?

Defining subgroups where FIT thresholds may need to be altered

35. i. We have defined the subgroups in the PROSPERO protocol as:

People taking medications or with conditions which increase the risk of gastrointestinal bleeding, such as antiplatelets (including aspirin, clopidogrel, prasugrel, dipryidamole and ticagrelor), anticoagulants (including warfarin, heparin, direct oral anticoagulants such as rivaroxaban, apixaban, edoxoban and dabigatran), proton pump inhibitors

People with blood disorders (e.g., haemoglobin variants/haemoglobinopathies such as beta thalassemia) that could affect the performance of the test

Are there any additions to this list or terms we should look out for?

Measurements and diagnostic test accuracy of different tests and analysers

36. Is it reasonable to assume that diagnostic test accuracy will differ for a given threshold according to test, and according to test and analyser combination?

    If so, is this due simply to different absolute values being reported by different tests (e.g. test X always reports higher values than test Y), or does the underlying diagnostic accuracy of the test and analyser combination also differ?

Survey results:

Seven advisors completed the questionnaire. The individual answers and mean values obtained from the survey that were used to inform the proportion of patients who follow each of the pathways and each of the imaging tests are presented in Table 83 to Table 87. The time to diagnosis associated with each of the pathways and patients who are missed by the imaging tests, estimated from the clinicians answers is presented in Table 88.

Table 83: Proportion of patients receiving each of the management pathways following FIT results, clinicians individual answers

	FIT <10						FIT <100						current care					
Expert	**Watch & Wait**	**repeat FIT**	**eRS**	**18WW**	**2WW**	**other**	**Watch & Wait**	**repeat FIT**	**eRS**	**18WW**	**2WW**	**others**	**Watch & Wait**	**repeat FIT**	**eRS**	**18WW**	**2WW**	**others**
1	40%	20%	-			40%	-	-	-	-	85%	15%	60%	-	-	-	-	40%
2	75%	-	-	small proportion		-	-	-	-	-	most	-	most (or 2nd FIT)	most (or W&W)	-	-	-	-
3	60%	15%	-	20%	5%	-	0%	0%	-	0%	100 %	-	0%	0%	-	0%	100 %	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	0%					should be referred	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	-	-	-	50%	50%?	-	-	-	-	-	-

Table 84: Proportion of patients receiving each imaging test as the first test during 2WW referral

	2WW (no age specified)					2WW (<60)					2WW (60-69)					2WW (70-79)					2WW (80+)				
Expert	**COL**	**CTC**	**capsule COL**	**CT**	**Discharge**	**COL**	**CTC**	**capsule COL**	**CT**	**Discharge**	**COL**	**CTC**	**capsule COL**	**CT**	**Discharge ***	**COL**	**CTC**	**capsule COL**	**CT**	**Discharge ***	**COL**	**CTC**	**capsule COL**	**CT**	**Discharge**
1	90%	10%	1%	2%	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	95%	3%	1%	2%	-	90%	5%	0%	5%	-	70%	20%	0%	10%	-	50%	20%	0%	30%	-
4	largest proportion	-	-	-	-	largest proportion	-	-	-	-	largest proportion	-	-	-	-	largest proportion	-	-	-	-	largest proportion	-	may have	may have	on basis of performance status or patient choice
5	-	-	0%	-	-	-	-	0%	-	-	-	-	0%	-	-	-	-	0%	-	-	-	-	0%	-	-
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	95%	-	1.50%	5%	ocasionally if frail or bening	90%	-	1.50%	5%	ocasionally if frail or bening	85%	-	rarely	10%	ocasionally if frail or bening	70%	0%	rarely	30%	ocasionally if frail or bening

Table 85: Proportion of patients receiving each imaging test as the first test during 18WW referral

	2WW (no age specified)					2WW (<60)					2WW (60-69)					2WW (70-79)					2WW (80+)				
Expert	COL	CTC	capsule COL	CT	Discharge	COL	CTC	capsule COL	CT	Discharge	COL	CTC	capsule COL	CT	Discharge	COL	CTC	capsule COL	CT	Discharge	COL	CTC	capsule COL	CT	Discharge
1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	95%	3%	0%	2%	-	90%	5%	0%	5%	-	70%	20%	0%	10%	-	50%	20%	0%	30%	-
4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5	largest propo rtion	-	-	-	-	largest propo rtion	-	-	-	-	largest propo rtion	-	-	-	-	largest propo rtion	-	-	-	-	largest propo rtion	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	90%	very few	tiny %	minority (weight loss and frial pts)	-	85%	very few	tiny %	minority (weight loss and frial pts)	-	70%	very few	tiny %	minority (weight loss and frial pts)	-	60%	very few	tiny %	minority (weight loss and frial pts)	-

Table 86: Proportion of patients receiving each of the management pathways following FIT results, estimated values by the EAG based on clinicians' answers

	Proportion of person receiving each of the safety netting options		
	Low intentsity safety netting	High intensity safety netting	Intermediate FIT score FU
2WW	5.0%	15.0%	85.0%
18WW	10.0%	25.0%	10.0%
Watch and wait	75.0%	40.0%	0.0%
repeat FIT	10.0%	20.0%	5.0%

Table 87: Proportion of patients receiving each main imaging test during their lower GI referral (2WW/18WW)

	2WW					**18WW**				
Age group	Overall population	<60 years old	60 - 69 years old	70 - 79 years old	80+ years old	Overall population	<60 years old	60 - 69 years old	70 - 79 years old	80+ years old
COL	90%	95%	90%	70%	60%	90%	95%	90%	70%	50%
CTC	7.50%	3%	5%	20%	10%	7.50%	3%	5%	20%	20%
Other investigations	2.50%	2%	5%	10%	30%	2.50%	2%	5%	10%	30%
total check	100%	100 %	100 %	100 %	100 %	100%	100 %	100 %	100 %	100 %

Table 88: Time to diagnosis by type of pathway and diagnostic result (weeks), based on clincians ansewers to questionaire

	Estimated time to diagnosis (weeks)			
	CRC, AAs, IBD - base-case	**CRC, AAs, IBD - scenario 1**	**CRC, AAs, IBD - scenario 2**	**non-underlying pathology**
2WW	2	2	3	0
18WW	27	18	54	0
repeat FIT	59	35	104	0
long delay (FN)	78	52	157	0

Appendix 12: Model estimation of the impact of additional time to diagnosis on colorectal cancer outcomes

April 2023
Dr Sophie Whyte
Laura Heathcote
Dr Chloe Thomas
Dr Olena Mandrik

1. Introduction

As time goes by from the onset of symptoms to their presentation in primary care, the disease may progress for individuals with symptomatic CRC. Additionally, the disease may also progress as time passes from the primary care presentation to receiving a diagnosis.

Patients presenting to their GP with suspected CRC may have various underlying conditions that could explain their symptoms, including non-cancerous conditions such as inflammatory bowel disease (IBD), diverticulitis, irritable bowel syndrome (IBS), or haemorrhoids. Symptoms of these conditions can overlap with CRC symptoms and include abdominal pain, rectal bleeding, changes in bowel habits, and weight loss. Moreover, patients may have adenomas, which are generally asymptomatic but can be diagnosed incidentally during investigations for suspected CRC and are clinically important because adenomas (particularly advanced adenomas (HRA)) have the potential to develop into CRC.

The NHS two-week wait (2WW) system aims to expedite the diagnosis and treatment of patients with suspected cancer, including colorectal cancer (CRC), by ensuring that they are seen by a specialist within two weeks of referral by their GP. According to the latest annual data from NHS England (2020-2021), 88.9% of patients with suspected colorectal cancer referred through the two-week wait system were seen by a specialist within two weeks of referral, indicating that the vast majority of patients with suspected CRC are able to access specialist care within the recommended timeframe. [278]

The 2WW system is important because it facilitates early diagnosis and treatment of cancer, which is crucial for improving survival rates and reducing the need for costly and invasive treatments. The same data [278] shows that only 50.6% of patients with suspected colorectal cancer referred through the two-week wait system started their first treatment within 62 days of referral, well below the target of 85%. The data for 2019-2020 [279] shows that 66.7% of patients started their first treatment within 62 days of referral, suggesting that the failure to meet the target is not wholly explained by the extraordinary circumstances of the COVID-19 pandemic.

Patients who do not meet the criteria for an urgent referral may be referred via the non-urgent cancer referral pathway, which aims to ensure that patients with suspected cancer receive a diagnosis or are given the all-clear within 18 weeks from referral by their GP. The 18-week target applies to all non-urgent referrals, including those for suspected cancer.

Evidence on the association between time to diagnosis and CRC outcomes is heterogeneous. A systematic review explored the association between shorter times to diagnosis and more favourable outcome and found that although many studies reported no associations, more studies reported a positive, rather than a negative, association. [REF Neal 2015].

The objective of this study was to develop a health economic model of CRC progression for symptomatic individuals associated with additional time to diagnosis, and to use the model to estimate the impact of additional time to diagnosis in terms of healthcare costs and health outcomes.

2. Methods

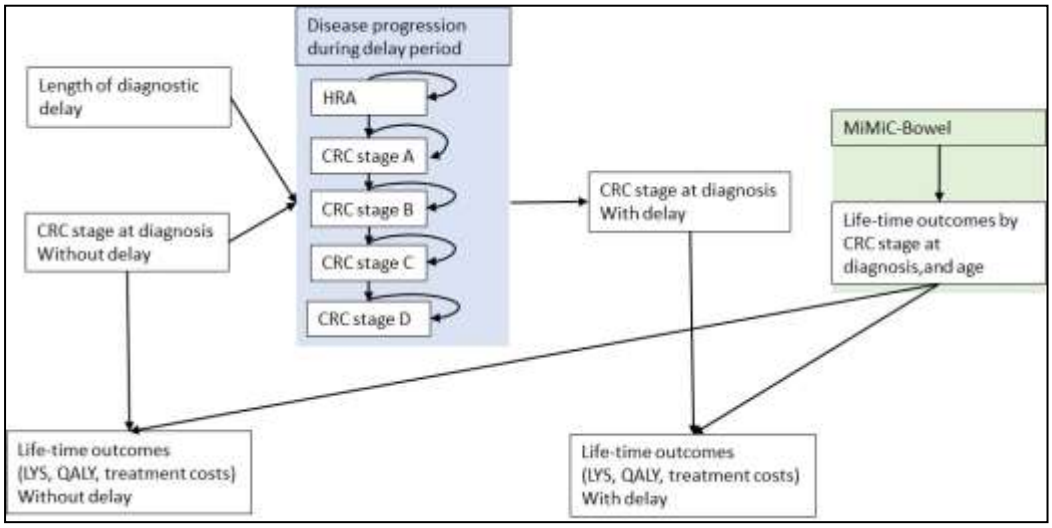
Model perspective

A lifetime horizon was adopted in this analysis to evaluate the long-term impact of additional time to diagnosis on healthcare costs and health outcomes for CRC and HRA. A discount rate of 3.5% was used to adjust for time preferences, in line with the NICE reference case. The analysis was conducted from the perspective of the UK National Health Service (NHS) and personal social services, which included all costs and benefits associated with healthcare services and social care interventions. Both direct and indirect costs associated with CRC diagnosis and treatment, including costs of diagnostic tests, healthcare contacts, hospitalisations, medications, and palliative care, were considered. Health benefits or disbenefits were measured in terms of life years gained (or lost) (LYG), and quality-adjusted life years (QALYs).

Model structure

A health economic model was used to estimate impact on patient outcomes of additional time to diagnosis. The model structure is illustrated in Figure 21. For patients with CRC, the impact of additional time to diagnosis is estimated by comparing the stage distribution of CRC at diagnosis without delay to the expected stage distribution of CRC at diagnosis with the delay. The change in stage distribution during the additional time to diagnosis represents disease progression during this time period. For patients with HRA, disease progression represented by the proportion of individuals who develop CRC during the additional time to diagnosis. These estimates of disease progression during the additional time to diagnosis are combined with estimates of the differential outcomes by disease stage to produce an overall estimate of the impact of additional time to diagnosis.

Figure 21: model diagram



Population

The model population reflects the population of patients in the 2WW system for suspected CRC in England [280]. All individuals in the model have either CRC or HRA.

The age distributions applied for both CRC and HRA are shown in Table 89.

Estimates for different age band were generated and these were combined to produce estimates for a cohort.

The stage distribution without additional time to diagnosis (i.e. at the start of the model) was assumed to correspond to the stage distribution for patients diagnosed via symptomatic or chance detection (i.e., not via screening or surveillance), as shown in Figure **22**. This stage distribution is a snapshot of the disease present at the point of data collection. Therefore, it reflects the stage distribution at the average time to diagnosis.

We note that chance detection may be associated with an earlier average stage at diagnosis than symptomatic presentation so this data may show an earlier stage distribution than is appropriate for a symptomatic population which is a minor limitation.

Figure 22: stage distribution of colorectal cancer in the 2WW population in England [281]

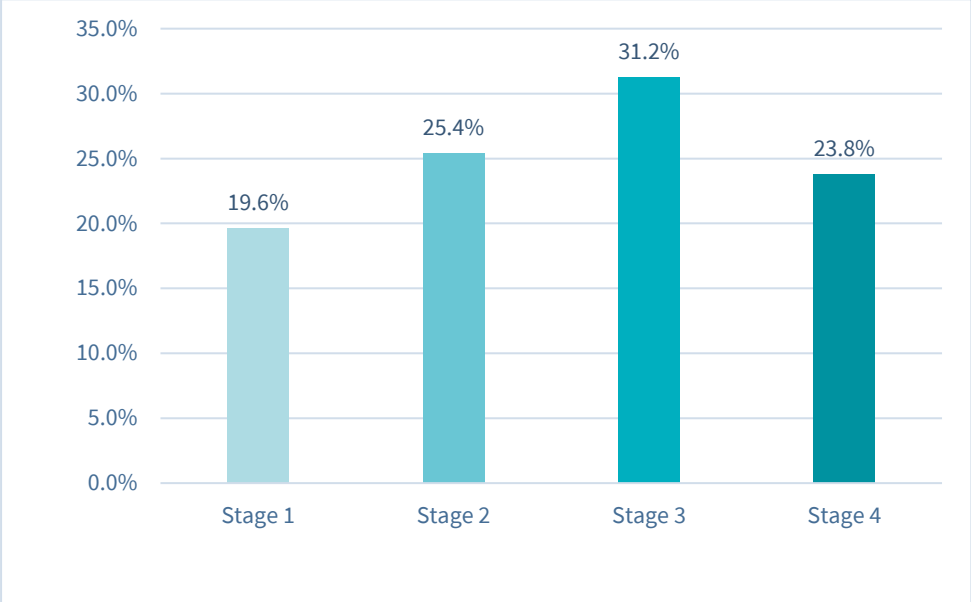


Table 89: age distributions

Age distribution used for CRC: CRC diagnosed via 2WW [280]		**Age distribution used for HRA:** 2WW referrals population [280]	
Frequency, N	%	Frequency, N	%
734	5%	49,251	13%
1,814	13%	63,396	17%
2,841	21%	85,690	23%
4,274	32%	104,062	28%
3,789	28%	73,564	20%
13,452	100%	375,963	100%

2.2 Modelling disease progression during delay period

Patients start the model in one of five health states: advanced adenoma (HRA), or CRC stage A, B, C or D. During the additional time to diagnosis, a proportion of patients will experience a stage shift, i.e. a proportion of patients with HRA will develop CRC stage A, and a proportion of patients in each CRC stage will advance to the next stage. The probability of transitioning depends on the length of additional time to diagnosis. It is assumed that patients can only make up to one transition within one year. It is assumed that all patients survive the delay period, i.e. there is no transition to "dead" in the model. To include deaths within the diagnosis delay period would require updating the delay progression component to depend on age which would add a

fair amount of complexity. This is a minor limitation of the methods but is not expected to have a significant impact on the results.

The transition probabilities in the model were taken from MiMiC Bowel [96]. MiMiC Bowel is a microsimulation model of CRC written in R, which includes a natural history model.

The calculations assume that persons can only make one state transition within a 1-year period. This assumption is consistent with the assumptions made within MiMiC-Bowel model. For predictions related to delays of >1 year, multiple transitions are included.

MiMiC-Bowel reports annual transition probabilities however to estimate transitions for shorter periods of time it is necessary to first convert transition probabilities into rates. The formula used was: rate, r=-ln [1-annual_trans_prob], then to estimate the transition probabilities relating to shorter time period the formula $p(t) = 1 - e^{-rt}$, where r is the rate and t is the time period was used. Note that this conversion formula has weaknesses and is most reliable for a model in which a person can experience only one type of event in a single cycle [282].

In MiMiC-Bowel, the preclinical patient population includes both asymptomatic and symptomatic patients, hence the preclinical disease progression probabilities therefore relate to both asymptomatic and symptomatic individuals. It is plausible that a wholly symptomatic population may experience faster disease progression and this has been explored within a sensitivity analysis.

Table 90: transition probabilities used within the model [96]

Transition	MiMiC-Bowel, transition probability (1 year)	MiMiC-Bowel, transition rate (1 year)
CRC A -> CRC B	0.293	0.347
CRC B -> CRC C	0.554	0.807
CRC C -> CRC D	0.350	0.431
HR->CRC A *	0.027	0.028
***Risk of progression is age dependent for HRA ->CRC but average transition rate for age of 62 has been used currently.**		

2.3 Lifetime outcomes for CRC and HRA

CRC

Lifetime outcomes for CRC without additional time to diagnosis were estimated using MiMiC Bowel.

The model was set up to best reflect current practice in CRC screening and diagnosis, i.e. individuals in the model were eligible for screening by FIT test at the age of 56. The model records diagnoses and outcomes separately for individuals diagnosed via screening or via symptomatic presentation. Only outcomes for individuals diagnosed symptomatically were used, as this best represents individuals in the NHS 2WW pathway.

The model was run for a population of 169,975 individuals. For each individual diagnosed symptomatically, the life years, QALYs, and healthcare costs from the point of diagnosis until death were recorded. These outcomes were then subdivided according to the age group and stage at diagnosis, and the mean outcomes per age and stage at diagnosis were calculated. Details on how these outcomes are estimated by MiMiC Bowel are reported in full in the relevant published model documentation [96].

As the costs in MiMiC Bowel correspond to 2018 prices, aggregate costs were inflated to the latest possible price year (2021) using NHSCII from the Unit costs of health and social care 2021 [117].

HRA

It is implicitly assumed that individuals diagnosed with HRA have them removed via polypectomy. It was assumed that such individuals have the same lifetime cost and health outcomes as the general population. It is possible that individuals with HRA would be expected to have higher lifetime costs and less favourable lifetime health outcomes, however a simplifying assumption was made as this was not anticipated to have a significant impact on model outcomes.

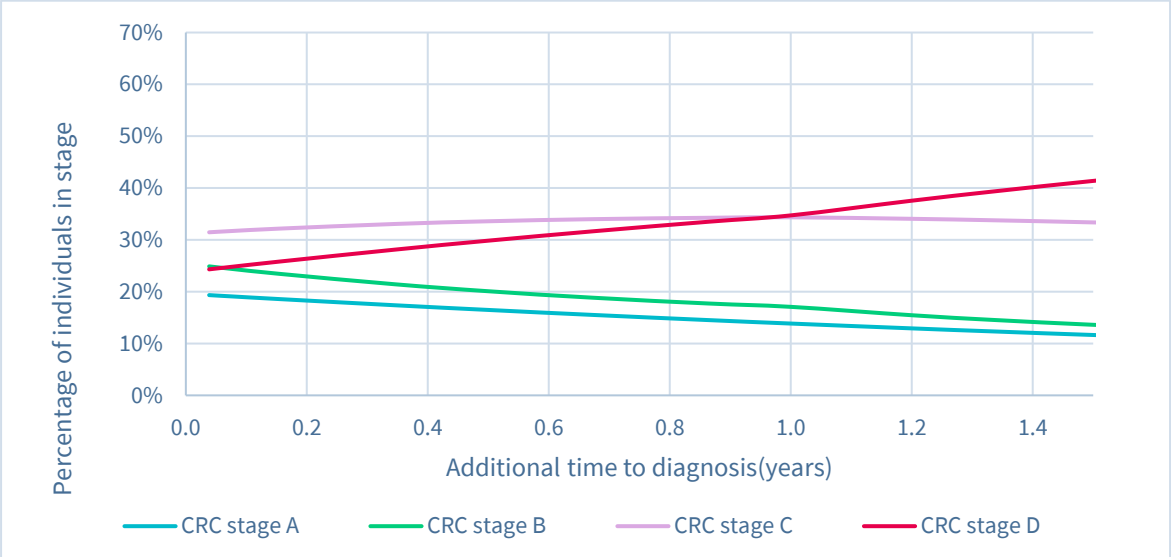
Life expectancy was taken from life tables [283] and age and sex-adjusted HRQoL was estimated using methods published by Ara et al [284].

3. Results

Model estimates of disease progression

Figure 23 shows the change in CRC stage distribution with increasing additional time to diagnosis. With longer additional time to diagnosis, more individuals progress to late stage (C or D) CRC, and fewer are diagnosed in early stages (A or B).

Figure 23: change in CRC stage distribution by additional time to diagnosis



Costs and health outcomes by age and stage

Table 91 shows expected outcomes by age and stage, as estimated by MiMiC Bowel. There is an inconsistency in the results in that outcomes are less favourable for HRA than for stage A CRC- this is likely due to uncertainty in the model outcomes and the overwhelmingly positive outcomes associated with an early diagnosis of CRC. Within CRC stages, later stage is associated with fewer LYs, as is older age at diagnosis.

Fewer lifetime QALYs are accrued by individuals with CRC than with HRA; and within CRC fewer QALYs are accrued by individuals diagnosed at later stages than at early stages. Within each stage, individuals in older age groups accrue fewer lifetime QALYs than those diagnosed in younger age groups. Expected QALY estimates are lower than the corresponding LY estimates, reflecting the impact of CRC and CRC treatment on HRQoL.

Lifetime treatment costs show a more complex pattern. Treatment costs for individuals with CRC are much higher than for individuals with HRA. Individuals diagnosed with stage D cancer have the lowest treatment costs (likely due to such individuals having much shorter life expectancy, and more likely to be offered only palliative treatment). The pattern across the other age groups and stages is influenced by the interactions between life expectancy and treatment options.

Table 91: expected discounted LYs, QALYs, and inflated treatment costs by age and stage at diagnosis

Age	Expected discounted lifetime LYs Diagnosed with HR adenomas	Diagnosed with CRC stage A	Diagnosed with CRC stage B	Diagnosed with CRC stage C	Diagnosed with CRC stage D
< 50	22.5	22.7	21.4	20.0	6.6
50-59	18.2	18.6	18.2	16.3	6.0
60-69	14.5	15.7	14.6	13.2	5.3
70-79	10.2	11.4	10.6	9.1	4.3
80+	5.8	7.3	7.0	6.1	4.0

Age	Expected discounted lifetime QALYs Diagnosed with HR adenomas	Diagnosed with CRC stage A	Diagnosed with CRC stage B	Diagnosed with CRC stage C	Diagnosed with CRC stage D
< 50	19.3	13.7	11.3	10.1	1.5
50-59	14.7	8.5	8.2	6.7	1.3
60-69	11.5	6.0	5.5	4.5	1.1
70-79	7.4	3.6	3.1	2.4	0.8
80+	3.4	1.9	1.6	1.3	0.7

Age	Expected discounted lifetime treatment costs (INFLATED TO 2021) Diagnosed with HR adenomas	Diagnosed with CRC stage A	Diagnosed with CRC stage B	Diagnosed with CRC stage C	Diagnosed with CRC stage D
< 50	£530	£ 14,621	£ 13,930	£ 19,762	£ 7,166
50-59	£537	£ 14,602	£ 14,800	£ 19,186	£ 5,536
60-69	£481	£ 15,972	£ 15,521	£ 16,662	£ 4,533
70-79	£355	£ 14,646	£ 13,303	£ 13,486	£ 3,215
80+	£87	£ 11,297	£ 10,317	£ 10,402	£ 2,091

Impact of additional time to diagnosis

Table 92 shows the estimated impact of additional time to diagnosis for individuals with CRC. Note that *additional* refers to beyond the current time to diagnosis on the 2WW pathway- time zero is current time to diagnosis. All results are incremental compared to this. With increasing additional time to diagnosis, health outcomes (LYs and QALYs) are worse. Treatment costs are also lower (due to more individuals being diagnosed in stage D which has lower treatment costs). However, at the willingness-to-pay (WTP) thresholds of £20,000 or £30,000 per QALY, the QALY loss outweighs the treatment cost savings, resulting in lower net monetary benefit (NMB) with increasing additional time to diagnosis.

Table 92: estimated outcomes by additional time to diagnosis for CRC (NMB = Net Monetary Benefit, WTP = Willingness To Pay). All outcomes are discounted at 3.5% per annum.

Additional time to diagnosis

months	**0.0**	**0.5**	**1.0**	**2.0**	**4.0**	**6.0**	**8.0**	**10.0**	**12.0**
weeks	**0.0**	**2.2**	**4.3**	**8.7**	**17.3**	**26.0**	**34.7**	**43.3**	**52.0**
Incremental values versus time zero									
NMB (WTP=£30k/QALY)	£0	-£573	-£1,133	-£2,220	-£4,748	-£6,591	-£8,494	-£10,414	-£11,649
NMB (WTP=£20k/QALY)	£0	-£366	-£724	-£1,418	-£3,031	-£4,205	-£5,415	-£6,636	-£7,419
LYs	0.00	-0.04	-0.08	-0.17	-0.35	-0.49	-0.64	-0.78	-0.88
QALYs	0.00	-0.02	-0.04	-0.08	-0.17	-0.24	-0.31	-0.38	-0.42
Treatment costs	£0	-£47	-£94	-£186	-£404	-£568	-£742	-£922	-£1,041
Absolute values									
LYs	10.03	9.98	9.94	9.86	9.67	9.53	9.39	9.24	9.15
QALYs	3.37	3.35	3.33	3.29	3.19	3.13	3.06	2.99	2.94
Treatment costs	£11,458	£11,410	£11,364	£11,272	£11,054	£10,890	£10,716	£10,536	£10,417

HRA

Table 93 shows the impact of additional time to diagnosis of HRA. With increasing time to diagnosis, more LYs are accrued. This is likely due to the inconsistencies described previously in the expected LYs between HRA and CRC stage A. However, with increasing additional time to diagnosis, fewer QALYs are accrued (reflecting the lower HRQoL with stage A CRC versus HRA). Treatment costs are also higher, reflecting the higher treatment costs for CRC versus HRA.

Table 93: impact of additional time to diagnosis of HRA (NMB = Net Monetary Benefit, WTP = Willingness To Pay). All outcomes are discounted at 3.5% per annum.

Additional time to diagnosis

months	**0.0**	**0.5**	**1.0**	**2.0**	**4.0**	**6.0**	**8.0**	**10.0**	**12.0**
weeks	**0.0**	**2.2**	**4.3**	**8.7**	**17.3**	**26.0**	**34.7**	**43.3**	**52.0**
Incremental values versus time zero									
NMB (WTP=£30k/QALY)	£0	-£156	-£312	-£623	-£1,397	-£2,014	-£2,704	-£3,468	-£4,000
NMB (WTP=£20k/QALY)	£0	-£109	-£218	-£435	-£976	-£1,406	-£1,889	-£2,422	-£2,793
LYs	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.02	0.03
QALYs	0.00	-0.00	-0.01	-0.02	-0.04	-0.06	-0.08	-0.10	-0.12
Treatment costs	£0	£15	£30	£59	£133	£192	£257	£330	£381
Absolute values									
LYs	13.27	13.28	13.28	13.28	13.28	13.29	13.29	13.30	13.30
QALYs	10.35	10.35	10.34	10.33	10.31	10.29	10.27	10.25	10.23
Treatment costs	£385	£400	£415	£444	£518	£577	£642	£715	£766

Appendix 13: Methods for pooling prevalence data from the EAG clinical review

Data on prevalence were pooled using a Bayesian Markov chain Monte Carlo (MCMC) approach based on a random effects meta-analysis.[285] The prevalence for the overall population and by different population type (NG12 high/medium- and DG30 low-risk groups) was analysed separately. The random effects model allowed for heterogeneity in the prevalence across studies within each population type. The model assumed that the log odds of study-specific prevalence are from a normal distribution, where the mean represents the overall population prevalence, and the variance represents heterogeneity among the studies.

Vague priors were assumed for all model parameters for the meta-analysis of CRC as the outcome using population type 1 studies because this was the analysis with the greatest number of included studies which allowed for appropriate estimation of the heterogeneity parameter. For all other meta-analyses, a vague prior was assumed for the mean and an informative prior generated using the posterior distribution of the heterogeneity parameter for the meta-analysis of CRC as the outcome using population type 1 studies was assumed for the heterogeneity parameter due to limited studies to inform the estimation of the heterogeneity parameter.

All analyses were conducted in the freely available software package WinBUGS and R, using the R2Winbugs interface package.[286] Convergence to the target posterior distributions was assessed using the Gelman-Rubin statistic.[41] The chains converged within 50,000 iterations so a burn-in of 50,000 iterations was used. A further 30,000 iterations of the Markov chain was retained to estimate parameters using one chain and thinning every 5 iterations. The absolute goodness of fit was checked by comparing the number of data points (which is the number of included studies) with the total residual deviance.

Appendix 14: Additional health economic results

Figure 24: iNMB for Intervention 1 assuming a threshold of £20,000 per QALY gained and high intensity safety netting

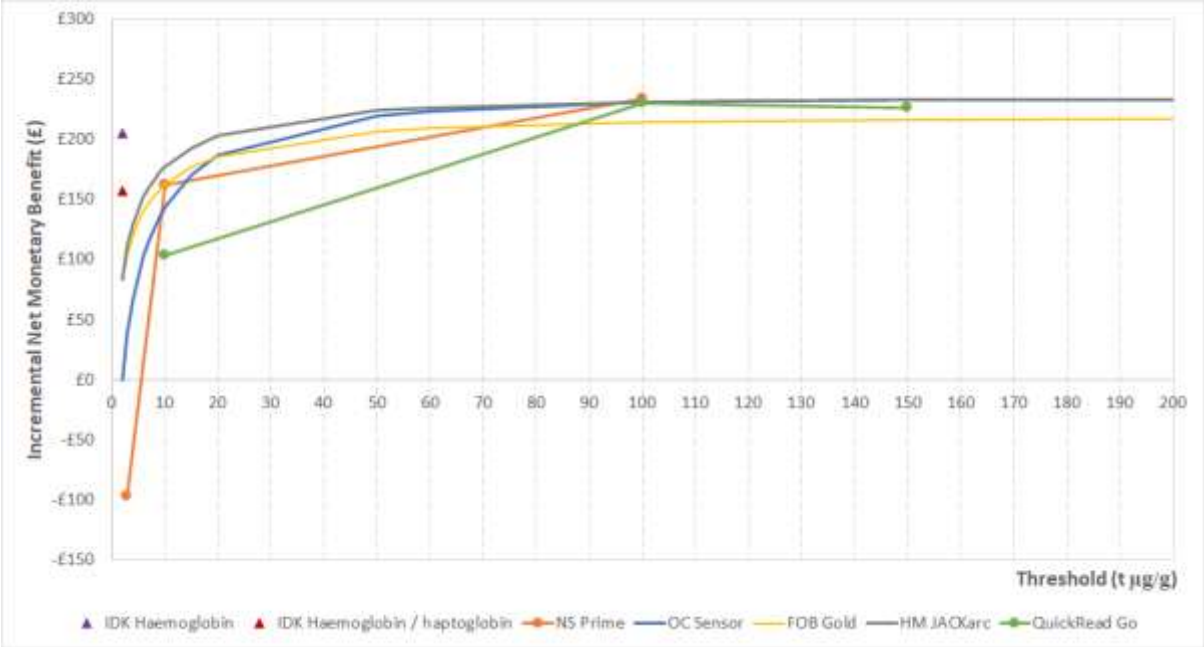


Figure 25: iNMB for Intervention 2 assuming a threshold of £20,000 per QALY gained and high intensity safety netting

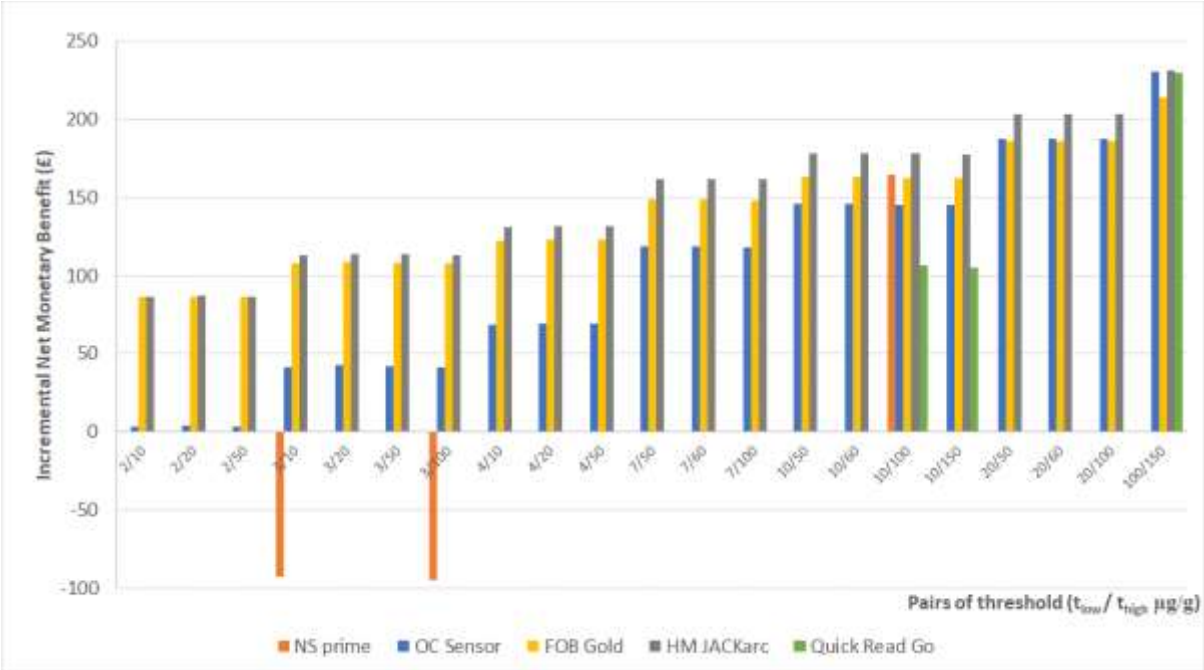


Figure 26: iNMB for Intervention 1 assuming a threshold of £30,000 per QALY gained and low intensity safety netting

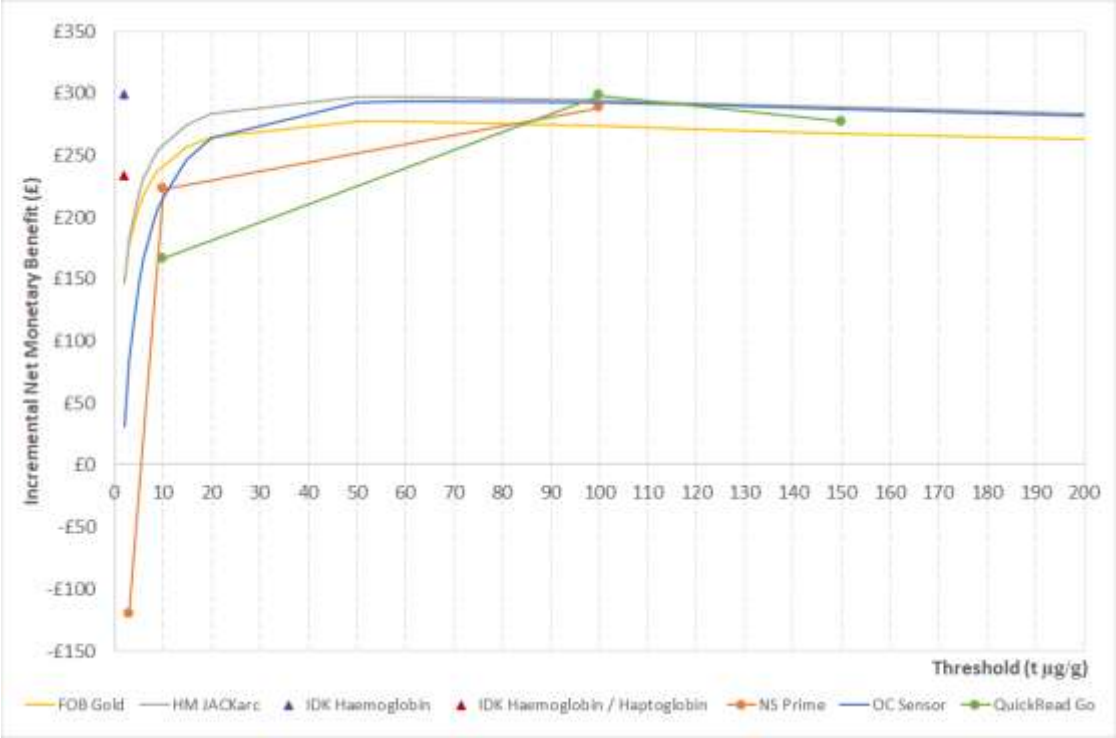


Figure 27: iNMB for Intervention 2 assuming a threshold of £30,000 per QALY gained and low intensity safety netting

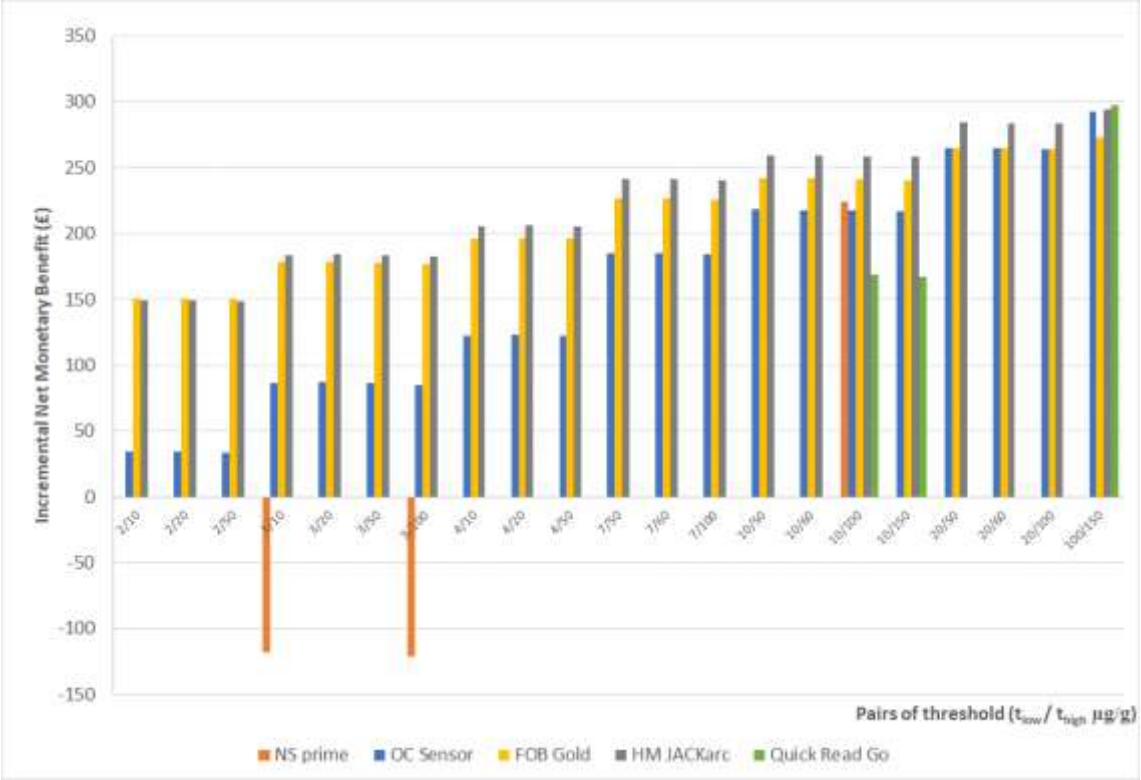


Figure 28: ***iNMB for Intervention 1 assuming a threshold of £30,000 per QALY gained and high intensity safety netting***

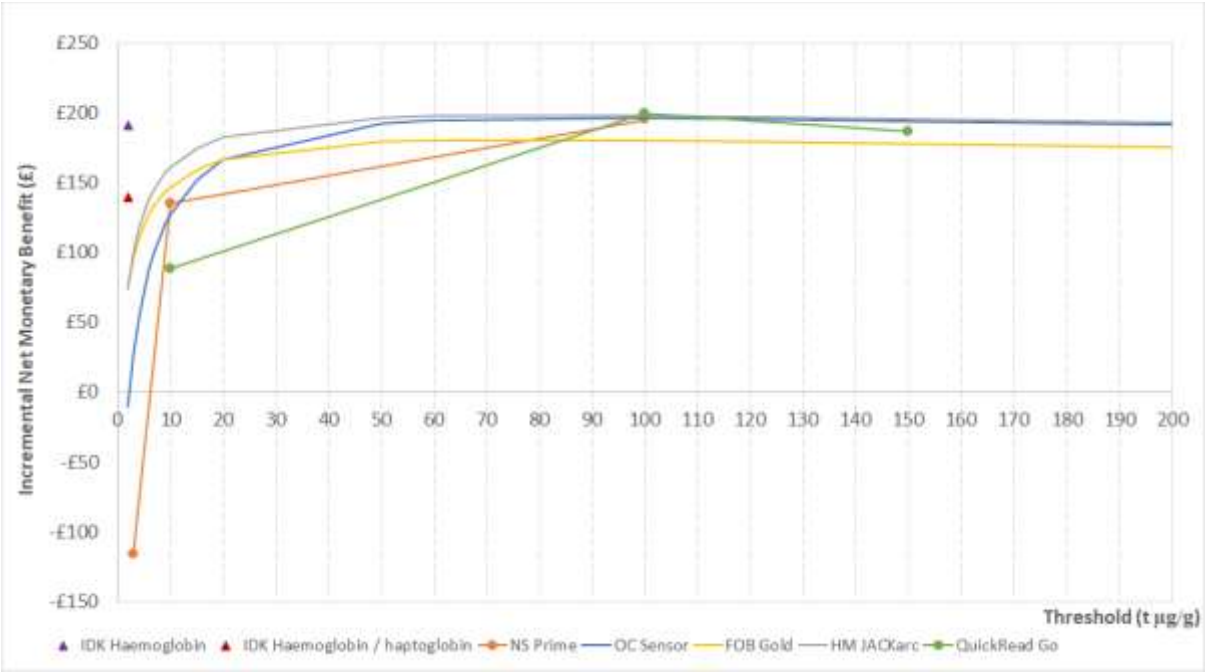
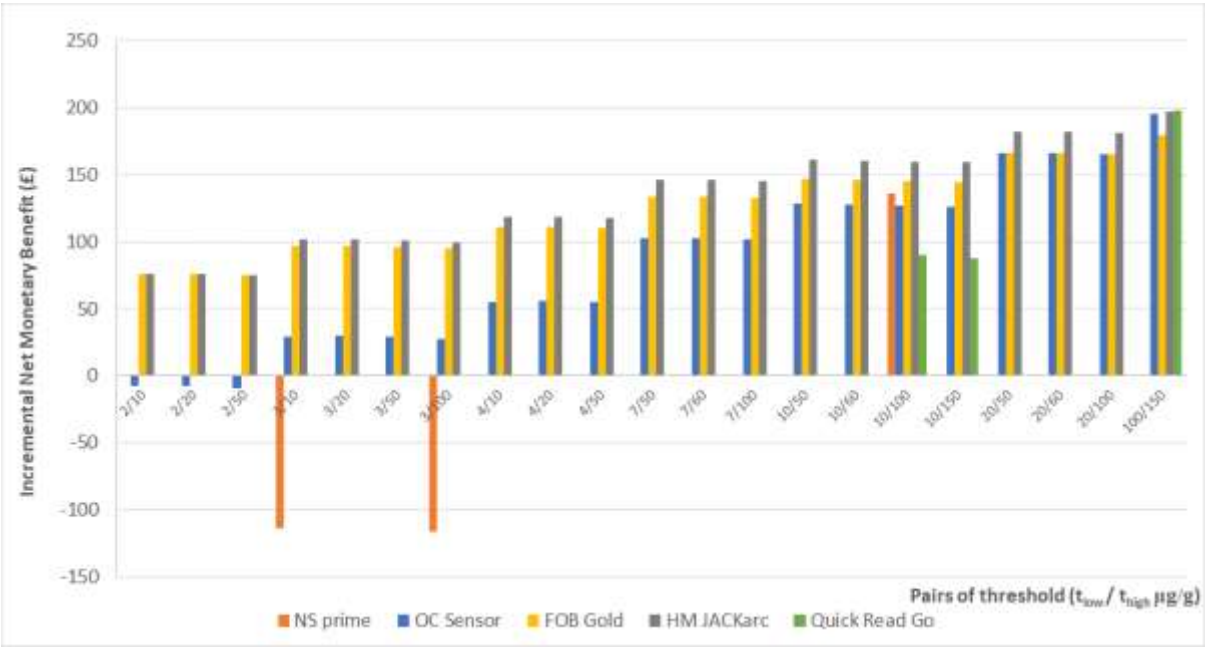


Figure 29: ***iNMB for Intervention 2 assuming a threshold of £30,000 per QALY gained and high intensity safety netting***



Tabulated results using high intensity safety netting

Table 94: Tabulated results for FOB Gold using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	**2**	**4**	**5**	**7**	**10**	**20**	**50**	**60**	**100**	**10**
LYs	14.167	14.167	14.167	14.167	14.167	14.166	14.165	14.165	14.164	14.168
QALYs	10.894	10.894	10.894	10.894	10.893	10.893	10.892	10.892	10.891	10.895
Costs (£)	3143	3102	3089	3071	3053	3021	2986	£2,980	2964	3,246
ICER (pairwise, vs Intervention 3)▪ (£)	109,643	127,482	129,408	129,234	125,789	115,702	97,078	93,293	83,271	-
NMB λ=20,000 (vs Int 3) (£)	84	121	132	147	162	186	206	£209	214	-
NMB λ=30,000 (vs Int 3) (£)	75	110	120	134	147	166	180	£181	180	-
Number of 2WW referrals (total)	0.450	0.560	0.024	0.002	0.001	3.751	0.022	0.249	0.004	0.681
Number of 18WW referrals (total)	0.162	0.178	0.182	0.189	0.196	0.207	0.219	0.221	0.226	0.094
Number of Repeat FITs (total)	0.129	0.142	0.146	0.151	0.157	0.166	0.175	0.177	0.181	0.075
Number of Watch and Wait (total) (total)	0.259	0.284	0.292	0.302	0.313	0.331	0.350	0.354	0.362	0.150
Number of COLs (total)	0.560	0.525	0.515	0.500	0.486	0.461	0.434	0.430	0.419	0.709
Reduction in number of referrals (total - 2WW + 18WW)	21.1%	26.0%	27.4%	29.5%	31.6%	35.1%	38.8%	39.4%	41.0%	-
Reduction in number of referrals (2WW only)	33.9%	41.8%	44.2%	47.6%	50.9%	56.5%	62.5%	63.5%	66.1%	-
Increase in number of referrals (18WW only)▪▪	72.6%	89.5%	94.5%	101.7%	108.8%	120.9%	133.6%	135.8%	141.3%	-
Reduction in number of COLs	21.0%	25.9%	27.3%	29.4%	31.5%	35.0%	38.7%	39.3%	40.9%	-
Mean time to diagnosis - CRC	1.899	1.991	2.034	2.112	2.219	2.510	3.106	3.257	3.737	1.303
Mean time to diagnosis - AAs	3.751	4.223	4.406	4.706	5.065	5.657	6.418	6.564	6.937	1.956
Mean time to diagnosis - IBD	2.317	2.606	2.717	2.902	3.121	3.609	4.312	4.453	4.838	1.684

▪Southwest quadrant ICER; ▪▪ Also the value for increased repeat FITs and increased number of watch and waits
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 95: Tabulated results for FOB Gold using two thresholds

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.167	14.167	14.167	14.167	14.167	14.167	14.167	14.167	14.166	14.166	14.166	14.166	14.168
QALYs	10.894	10.894	10.894	10.894	10.894	10.894	10.893	10.893	10.893	10.893	10.893	10.893	10.895
Costs (£)	3139	3138	3137	3100	3098	3097	3067	3066	3050	3049	3019	3018	3246
ICER (pairwise, vs Intervention 3)□ (£)	106,058	102,673	96,536	124,680	121,369	115,243	120,079	115,131	118,951	114,619	112,467	109,257	-
NMB λ=20,000 (vs Int 3) (£)	86	87	86	122	123	123	149	148	163	163	186	186	-
NMB λ=30,000 (vs Int 3) (£)	76	76	75	111	111	110	134	133	147	145	166	165	-
Number of 2WW referrals (total)	0.430	0.423	0.416	0.385	0.379	0.371	0.339	0.335	0.321	0.316	0.289	0.285	0.681
Number of 18WW referrals (total)	0.175	0.180	0.185	0.185	0.189	0.194	0.201	0.204	0.205	0.208	0.212	0.215	0.094
Number of Repeat FITs (total)	0.136	0.138	0.141	0.146	0.148	0.150	0.157	0.159	0.161	0.163	0.168	0.169	0.075
Number of Watch and Wait (total) (total)	0.259	0.259	0.259	0.284	0.284	0.284	0.302	0.302	0.313	0.313	0.331	0.331	0.150
Number of COLs (total)	0.554	0.552	0.549	0.522	0.520	0.518	0.495	0.493	0.481	0.480	0.459	0.457	0.709
Reduction in number of referrals (total - 2WW + 18WW)	21.9%	22.2%	22.5%	26.4%	26.7%	27.0%	30.3%	30.5%	32.2%	32.4%	35.4%	35.6%	-
Reduction in number of referrals (2WW only)	36.9%	37.9%	39.0%	43.4%	44.4%	45.5%	50.2%	50.8%	52.9%	53.6%	57.6%	58.2%	-
Increase in number of referrals (18WW only)	87.1%	91.9%	97.0%	97.2%	102.0%	107.1%	114.5%	117.5%	118.8%	121.8%	126.0%	129.0%	-
Increase in number of repeat FITs	81.7%	84.7%	87.9%	94.3%	97.3%	100.5%	109.7%	111.6%	115.0%	116.9%	124.1%	126.0%	-
Increase in number of watch and waits	72.6%	72.6%	72.6%	89.5%	89.5%	89.5%	101.7%	101.7%	108.8%	108.8%	120.9%	120.9%	-

Reduction in number of COLs	21.9%	22.2%	22.5%	26.3%	26.6%	26.9%	30.2%	30.4%	32.1%	32.3%	35.3%	35.5%	-
Mean time to diagnosis - CRC	1.934	1.966	2.031	2.015	2.046	2.108	2.213	2.278	2.308	2.371	2.568	2.629	1.303
Mean time to diagnosis - AAs	3.892	3.956	4.039	4.309	4.371	4.450	4.878	4.931	5.199	5.251	5.730	5.780	1.956
Mean time to diagnosis - IBD	2.405	2.458	2.535	2.660	2.712	2.786	3.046	3.100	3.241	3.294	3.678	3.730	1.684

▪Southwest quadrant ICER

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 96: Tabulated results for HM-JACKarc using one threshold

	Int 1: FIT 1 threshold (µg/g)									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.167	14.167	14.167	14.167	14.166	14.166	14.165	14.165	14.164	14.168
QALYs	10.894	10.894	10.894	10.893	10.893	10.893	10.892	10.892	10.892	10.895
Costs (£)	3142	3093	3078	3057	3036	3002	2967	2962	2947	3246
ICER (pairwise, vs Intervention 3)▫ (£)	103,355	127,434	130,556	131,884	129,391	120,170	102,198	98,501	88,769	-
NMB λ=20,000 (vs Int 3) (£)	84	129	142	160	177	203	224	226	231	-
NMB λ=30,000 (vs Int 3) (£)	74	117	129	146	161	183	197	198	198	-
Number of 2WW referrals (total)	0.451	0.387	0.368	0.342	0.317	0.276	0.237	0.230	0.215	0.681
Number of 18WW referrals (total)	0.162	0.180	0.186	0.194	0.201	0.213	0.225	0.226	0.231	0.094
Number of Repeat FITs (total)	0.129	0.144	0.149	0.155	0.161	0.170	0.180	0.181	0.185	0.075
Number of Watch and Wait (total) (total)	0.258	0.289	0.297	0.310	0.321	0.340	0.359	0.362	0.370	0.150
Number of COLs (total)	0.560	0.519	0.507	0.490	0.474	0.448	0.422	0.418	0.408	0.709
Reduction in number of referrals (total - 2WW + 18WW)	21.0%	26.8%	28.5%	30.9%	33.2%	36.9%	40.5%	41.1%	42.5%	2.6%
Reduction in number of referrals (2WW only)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	0.0%
Increase in number of referrals (18WW only)▫▫	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	20.9%	26.8%	28.4%	30.8%	33.1%	36.8%	40.4%	41.0%	42.4%	-
Mean time to diagnosis - CRC	1.973	2.077	2.123	2.208	2.320	2.611	3.167	3.302	3.720	1.303
Mean time to diagnosis - AAs	3.752	4.215	4.395	4.692	5.047	5.635	6.394	6.540	6.914	1.956
Mean time to diagnosis - IBD	2.318	2.599	2.707	2.889	3.106	3.590	4.291	4.432	4.818	1.684

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 97: Tabulated results for HM-JACKarc using two thresholds

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.167	14.167	14.167	14.167	14.167	14.167	14.166	14.166	14.166	14.166	14.166	14.166	14.168
QALYs	10.894	10.894	10.894	10.894	10.894	10.894	10.893	10.893	10.893	10.893	10.893	10.893	10.895
Costs (£)	3138	3136	3135	3090	3089	3087	3053	3052	3033	3032	3000	2999	3,246
ICER (pairwise, vs Intervention 3)▪ (£)	100,733	97,831	92,645	124,982	121,963	116,493	123,507	119,156	123,091	119,257	117,167	114,291	-
NMB λ=20,000 (vs Int 3) (£)	87	87	87	131	131	131	162	161	178	178	203	203	-
NMB λ=30,000 (vs Int 3) (£)	76	76	75	118	118	118	146	145	161	160	182	182	-
Number of 2WW referrals (total)	0.427	0.420	0.413	0.374	0.367	0.360	0.323	0.320	0.303	0.299	0.269	0.266	0.681
Number of 18WW referrals (total)	0.177	0.182	0.187	0.189	0.193	0.198	0.206	0.209	0.210	0.213	0.217	0.220	0.094
Number of Repeat FITs (total)	0.137	0.139	0.142	0.148	0.151	0.153	0.161	0.162	0.165	0.167	0.173	0.174	0.075
Number of Watch and Wait (total) (total)	0.258	0.258	0.258	0.289	0.289	0.289	0.310	0.310	0.321	0.321	0.340	0.340	0.150
Number of COLs (total)	0.553	0.551	0.549	0.515	0.513	0.511	0.485	0.483	0.470	0.469	0.446	0.445	0.709
Reduction in number of referrals (total - 2WW + 18WW)	22.0%	22.3%	22.6%	27.4%	27.7%	28.0%	31.7%	31.9%	33.8%	34.0%	37.2%	37.3%	-
Reduction in number of referrals (2WW only)	37.3%	38.3%	39.4%	45.1%	46.1%	47.1%	52.5%	53.1%	55.6%	56.1%	60.5%	61.0%	-
Increase in number of referrals (18WW only)	89.2%	94.3%	99.3%	101.3%	106.3%	111.4%	119.8%	122.5%	124.5%	127.2%	132.1%	134.9%	-
Increase in number of repeat FITs	82.9%	86.0%	89.2%	98.0%	101.2%	104.3%	114.8%	116.5%	120.7%	122.4%	130.2%	132.0%	-
Increase in number of watch and waits	72.4%	72.4%	72.4%	92.5%	92.5%	92.5%	106.5%	106.5%	114.4%	114.4%	127.1%	127.1%	-
Reduction in number of COLs	21.9%	22.2%	22.5%	27.3%	27.6%	27.9%	31.6%	31.8%	33.7%	33.9%	37.1%	37.3%	-
Mean time to diagnosis - CRC	2.012	2.044	2.104	2.102	2.133	2.191	2.305	2.361	2.404	2.458	2.664	2.717	1.303
Mean time to diagnosis - AAs	3.891	3.955	4.038	4.300	4.361	4.440	4.861	4.914	5.179	5.230	5.707	5.757	1.956
Mean time to diagnosis - IBD	2.404	2.457	2.534	2.652	2.703	2.776	3.031	3.084	3.223	3.276	3.658	3.709	1.684

▪Southwest quadrant ICER

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold

Table 98: Tabulated results for IDK Haemoglobin using one threshold

	Int 1: FIT 1threshold	Int 3: DG30&NG12
t (µg/g)	**2**	**10**
LYs	14.166	14.168
QALYs	10.893	10.895
Costs (£)	£3,012	£3,246
ICER (pairwise, vs Intervention 3)▪ (£)	£167,120	-
NMB λ=20,000 (vs Int 3) (£)	£205	-
NMB λ=30,000 (vs Int 3) (£)	£191	-
Number of 2WW referrals (total)	0.285	0.681
Number of 18WW referrals (total)	0.210	0.094
Number of Repeat FITs (total)	0.168	0.075
Number of Watch and Wait (total) (total)	0.336	0.150
Number of COLs (total)	0.454	0.709
Reduction in number of referrals (total - 2WW + 18WW)	36.1%	-
Reduction in number of referrals (2WW only)	58.1%	-
Increase in number of referrals (18WW only) ▪▪	124.4%	-
Reduction in number of COLs	35.9%	-
Mean time to diagnosis - CRC	2.466	1.303
Mean time to diagnosis - AAs	3.631	1.956
Mean time to diagnosis - IBD	2.199	1.684

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 99: Tabulated results for IDK Haemoglobin/Hapto using one threshold

	Int 1: FIT 1threshold	**Int 3: DG30&NG12**
t (µg/g)	**2**	**10**
LYs	14.165	14.168
QALYs	10.893	10.895
Costs (£)	£3,054	£3,246
ICER (pairwise, vs Intervention 3)▪ (£)	£111,534	-
NMB λ=20,000 (vs Int 3) (£)	£157	-
NMB λ=30,000 (vs Int 3) (£)	£140	-
Number of 2WW referrals (total)	0.336	0.681
Number of 18WW referrals (total)	0.195	0.094
Number of Repeat FITs (total)	0.156	0.075
Number of Watch and Wait (total) (total)	0.312	0.150
Number of COLs (total)	0.487	0.709
Reduction in number of referrals (total - 2WW + 18WW)	31.4%	2.6%
Reduction in number of referrals (2WW only)	50.6%	0.0%
Increase in number of referrals (18WW only) ▪▪	108.3%	-
Reduction in number of COLs	31.3%	-
Mean time to diagnosis - CRC	2.791	1.303
Mean time to diagnosis - AAs	3.668	1.956
Mean time to diagnosis - IBD	2.236	1.684

▪Southwest quadrant ICER; ▫▫ Also the value for increased repeat FITs and increased number of watch and waits
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 100: Tabulated results for NS Prime

	Int 1: FIT 1 threshold			Int 2: FIT 2 thresholds			Int 3: DG30&NG12
Threshold - t or t_{low}/t_{high} (µg/g)	**3**	**10**	**100**	**3/10**	**3/100**	**10/100**	**10**
LYs	14.165	14.164	14.163	14.165	14.165	14.164	14.168
QALYs	10.893	10.892	10.891	10.893	10.893	10.892	10.895
Costs (£)	3,303	3,029	2,933	3,297	3,296	3,025	£3,246
ICER (pairwise, vs Intervention 3)[□] (£)	Dominated	79,700	79,566	Dominated	Dominated	77,913	-
NMB λ=20,000 (vs Int 3) (£)	-97	162	234	-93	-95	164	-
NMB λ=30,000 (vs Int 3) (£)	-116	135	195	-114	-117	136	-
Number of 2WW referrals (total)	0.683	0.307	0.196	0.616	0.597	0.288	0.681
Number of 18WW referrals (total)	0.093	0.204	0.237	0.138	0.151	0.217	0.094
Number of Repeat FITs (total)	0.075	0.163	0.189	0.097	0.103	0.170	0.075
Number of Watch and Wait (total) (total)	0.149	0.326	0.379	0.149	0.149	0.326	0.150
Number of COLs (total)	0.709	0.468	0.396	0.689	0.683	0.462	0.709
Reduction in number of referrals (total - 2WW + 18WW)	-0.1%	34.1%	44.2%	2.7%	3.6%	34.9%	-
Reduction in number of referrals (2WW)	-0.2%	54.9%	71.3%	9.5%	12.4%	57.8%	-
Increase in number of referrals (18WW)	-0.4%	117.4%	152.5%	46.7%	60.8%	131.4%	-
Increase in number of repeat FITs	-0.4%	117.4%	152.5%	29.1%	37.8%	126.2%	-
Increase in number of watch and waits	-0.4%	117.4%	152.5%	-0.4%	-0.4%	117.4%	-
Reduction in number of COLs	-0.0%	34.0%	44.2%	2.8%	3.6%	34.8%	-
Mean time to diagnosis - CRC	2.828	3.515	4.348	2.922	3.031	3.598	1.303
Mean time to diagnosis - AAs	4.220	5.038	6.888	4.332	4.568	5.218	1.956
Mean time to diagnosis - IBD	2.676	3.097	4.795	2.738	2.954	3.265	1.684

[▪]*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 101: Tabulated results for OC Sensor using one threshold

	Int 1: FIT 1 threshold (µg/g)									Int 3 : DG30& NG12
t (µg/g)	**2**	**4**	**5**	**7**	**10**	**20**	**50**	**60**	**100**	**10**
LYs	14.167	14.167	14.167	14.167	14.166	14.166	14.165	14.164	14.164	14.168
QALYs	10.894	10.894	10.894	10.893	10.893	10.893	10.892	10.892	10.891	10.895
Costs (£)	3224	3155	3133	3101	3069	3018	2970	2963	2946	3246
ICER (pairwise, vs Intervention 3)▪ (£)	19,937	72,611	85,184	99,204	107,611	111,322	99,366	96,031	86,711	-
NMB λ=20,000 (vs Int 3) (£)	0	66	86	116	143	187	220	224	231	-
NMB λ=30,000 (vs Int 3) (£)	-11	53	73	101	127	166	192	194	196	-
Number of 2WW referrals (total)	0.561	0.465	0.435	0.394	0.354	0.292	0.237	0.229	0.210	0.681
Number of 18WW referrals (total)	0.129	0.157	0.166	0.178	0.190	0.208	0.224	0.227	0.232	0.094
Number of Repeat FITs (total)	0.103	0.126	0.133	0.143	0.152	0.167	0.180	0.181	0.186	0.075
Number of Watch and Wait (total) (total)	0.206	0.252	0.266	0.285	0.304	0.333	0.359	0.363	0.372	0.150
Number of COLs (total)	0.631	0.569	0.550	0.523	0.498	0.458	0.422	0.417	0.405	0.709
Reduction in number of referrals (total - 2WW + 18WW)	10.9%	19.7%	22.4%	26.2%	29.8%	35.4%	40.5%	41.2%	42.9%	-
Reduction in number of referrals (2WW only)	17.6%	31.8%	36.1%	42.2%	48.0%	57.1%	65.2%	66.4%	69.1%	-
Increase in number of referrals (18WW only)▪▪	37.7%	68.1%	77.3%	90.3%	102.7%	122.2%	139.5%	142.0%	147.9%	-
Reduction in number of COLs	10.9%	19.7%	22.4%	26.1%	29.7%	35.4%	40.4%	41.1%	42.8%	-
Mean time to diagnosis - CRC	2.008	2.094	2.136	2.215	2.324	2.621	3.219	3.368	3.827	1.303
Mean time to diagnosis - AAs	3.833	4.279	4.453	4.739	5.084	5.653	6.394	6.538	6.908	1.956
Mean time to diagnosis - IBD	2.397	2.659	2.761	2.932	3.138	3.606	4.291	4.431	4.813	1.684

▪Southwest quadrant ICER; ▪▪ Also the value for increased repeat FITs and increased number of watch and waits

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 102: Tabulated results for OCSensor using two thresholds

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.167	14.167	14.167	14.167	14.167	14.167	14.166	14.166	14.166	14.166	14.166	14.166	14.168
QALYs	10.894	10.894	10.894	10.894	10.894	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.895
Costs (£)	3,219	3,218	3,216	3,152	3,150	3,148	3,095	3,094	3,065	3,064	3,016	3,015	3246
ICER (pairwise, vs Intervention 3)▪ (£)	23,029	23,333	22,832	72,762	71,716	68,852	94,118	90,657	103,073	99,702	108,718	105,930	-
NMB λ=20,000 (vs Int 3) (£)	3	4	4	68	69	69	119	118	146	145	188	187	-
NMB λ=30,000 (vs Int 3) (£)	-8	-8	-9	55	56	55	103	101	128	127	166	166	-
Number of 2WW referrals (total)	0.525	0.514	0.504	0.445	0.434	0.424	0.366	0.361	0.333	0.329	0.282	0.278	0.681
Number of 18WW referrals (total)	0.153	0.161	0.167	0.170	0.178	0.184	0.197	0.200	0.204	0.207	0.215	0.218	0.094
Number of Repeat FITs (total)	0.115	0.119	0.122	0.132	0.136	0.139	0.152	0.153	0.159	0.160	0.170	0.171	0.075
Number of Watch and Wait (total) (total)	0.206	0.206	0.206	0.252	0.252	0.252	0.285	0.285	0.304	0.304	0.333	0.333	0.150
Number of COLs (total)	0.620	0.617	0.614	0.563	0.560	0.557	0.515	0.514	0.492	0.490	0.455	0.454	0.709
Reduction in number of referrals (total - 2WW + 18WW)	12.5%	13.0%	13.4%	20.6%	21.1%	21.5%	27.4%	27.6%	30.7%	30.9%	35.9%	36.1%	-
Reduction in number of referrals (2WW only)	23.0%	24.6%	26.0%	34.7%	36.3%	37.7%	46.3%	47.0%	51.1%	51.8%	58.5%	59.2%	-
Increase in number of referrals (18WW only)	63.7%	71.5%	78.4%	81.9%	89.7%	96.6%	110.0%	113.3%	117.4%	120.8%	129.1%	132.4%	-
Increase in number of repeat FITs	53.9%	58.8%	63.1%	76.7%	81.6%	85.9%	102.6%	104.7%	111.9%	114.0%	126.5%	128.6%	-
Increase in number of watch and waits	37.7%	37.7%	37.7%	68.1%	68.1%	68.1%	90.3%	90.3%	102.7%	102.7%	122.2%	122.2%	-
Reduction in number of COLs	12.5%	13.0%	13.4%	20.5%	21.0%	21.4%	27.3%	27.5%	30.6%	30.8%	35.8%	36.0%	-
Mean time to diagnosis - CRC	2.048	2.083	2.154	2.120	2.153	2.219	2.321	2.385	2.415	2.477	2.679	2.739	1.303
Mean time to diagnosis - AAs	3.979	4.046	4.135	4.367	4.430	4.512	4.912	4.966	5.217	5.269	5.724	5.774	1.956
Mean time to diagnosis - IBD	2.486	2.542	2.623	2.713	2.765	2.841	3.075	3.131	3.256	3.310	3.673	3.724	1.684

▪*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 103: Tabulated results for QuikRead go

	Int 1: FIT 1 threshold			Int 2: FIT 2 thresholds			Int 3: DG30& NG12
t (µg/g)	**10**	**100**	**150**	**10/100**	**10/150**	**100/150**	**10**
LYs	14.167	14.164	14.163	14.166	14.166	14.164	14.168
QALYs	10.893	10.892	10.891	10.893	10.893	10.892	10.895
Costs (£)	3,112	2,953	2,939	3,106	3,105	2,953	3,246
ICER (pairwise, vs Intervention 3)□ (£)	88,897	94,197	76,466	83,756	79,595	91,998	-
NMB λ=20,000 (vs Int 3) (£)	103	230	227	107	105	229	-
NMB λ=30,000 (vs Int 3) (£)	88	199	187	90	88	198	-
Number of 2WW referrals (total)	0.408	0.218	0.204	0.374	0.372	0.215	0.681
Number of 18WW referrals (total)	0.174	0.230	0.234	0.197	0.198	0.232	0.094
Number of Repeat FITs (total)	0.139	0.184	0.187	0.151	0.151	0.185	0.075
Number of Watch and Wait (total)	0.279	0.368	0.375	0.279	0.279	0.368	0.150
Number of COLs (total)	0.532	0.410	0.401	0.522	0.521	0.409	0.709
Reduction in number of referrals (total - 2WW + 18WW)	24.9%	42.2%	43.5%	26.4%	26.5%	42.3%	-
Reduction in number of referrals (2WW only)	40.2%	68.0%	70.1%	45.1%	45.4%	68.4%	-
Increase in number of referrals (18WW only)	85.9%	145.5%	149.9%	109.7%	111.5%	147.3%	-
Increase in number of repeat FITs	85.9%	145.5%	149.9%	100.8%	101.9%	146.6%	-
Increase in number of watch and waits	85.9%	145.5%	149.9%	85.9%	85.9%	145.5%	-
Reduction in number of COLs	24.9%	42.1%	43.4%	26.3%	26.4%	42.3%	-
Mean time to diagnosis - CRC	2.148	3.436	4.356	2.285	2.382	3.521	1.303
Mean time to diagnosis - AAs	5.137	6.919	7.172	5.327	5.354	6.942	1.956
Mean time to diagnosis - IBD	3.185	4.822	5.116	3.360	3.392	4.849	1.684

▪*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – t

NICE National Institute for Health and Care Excellence

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

Section A: External Assessment Report - Comments

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Alpha Laboratories	1	23	3.3.3.2	The manufacturer name for the HM-JACKarc is now Minaris Medical Co., Ltd not *Hitachi Chemical Diagnostics Systems Ltd*.	Thank you, updated.
Alpha Laboratories	2	26	Table 1	Limit of detection for HM-JACKarc ~~0.6 µg/g~~ 2 µg/g Limit of quantitation for HM-JACKarc ~~1.25 µg/g~~ 7 µg/g	Thank you, updated.
Alpha Laboratories	3	195	Table 48	Historic costing: ~~£2.31~~ Current costing: £4.10	Thank you, the cost is now updated and the results provided as an addendum to the main EAG report.
Mast Group Ltd	4	8	2.4.1	In the paragraph describing test failure rates we propose that definition of 'failure' should be included here. Process errors/ pre-analytical errors e.g. non dated sample bottles, incorrect sample container used etc should be presented as separate to analytical errors.	There was insufficient time during the course of this assessment to report this level of detail and some studies did not report a definition. A sentence has been added to the discussion (6.2.1.1 and 6.2.1.2) and to the implications for service provision and generalisability sections (6.4 and 6.5). We have not changed section 2.4.1 as this reflects the scope issued by NICE.
Mast Group Ltd	5	23	3.3.3.4	Mast submitted information about a new analyser OC-Sensor CERES in the NICE DAP50 request for information update submitted in August 2022 to a Mr Jacob Grant. Is this mentioned in the redacted statement in 3.3.3.4? Performance is equivalent to the OC-Sensor PLEDIA.	Thank you for highlighting this data. We have added a critique of the comparisons provided to the clinical review, sections 4.3.2 and 6.2.1.1.

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
Mast Group Ltd	6	33	3.3.8.9	"A proportion of patients do not return their FIT tests. Based on the systematic review conducted for DG30, this ranged from 41% to 98% for OC Sensor, and 56%-66% for HM-JACKarc" We believe this is a transcription error. Uptake rates shown in the DG30 systemic review by Westwood et al states: "Reported uptake rates for the OCSensor studies included in our review varied widely, ranging from 41% (in a study where patients were sent an invitation to participate along with their referral letter [32]) to 98% (in a study where patients were given the specimen collection device at their initial consultation with a gastroenterologist [29])' Therefore figures presented demonstrate uptake not non-return rates and the sentence should be corrected.	Thank you for your comment, we have updated the report as follows: "A proportion of patients do not return their FIT tests. Based on the systematic review conducted for DG30, FIT was returned by 41% (in a study where patients were sent an invitation to participate along with their referral letter) to 98% (in a study where patients were given the specimen collection device at their initial consultation with a gastroenterologist) for patients using OC Sensor, and 56%-66% patients using HM-JACKarc. This was to be taken into account within the project.
Mast Group Ltd	7	33	3.3.8.9	We suggest a comment regarding factors affecting uptake should be included here as different distribution options are employed in England and have an impact on the uptake rates for FIT. For example (and this represents is just a few known routes), some regions offer a postal return route, some use assembled kits to be distributed by the GP, some offer a direct to patient route where kits are sent to patients' homes via royal mail. Other factors	We have not changed section 3.3.8.9 as this section is based on the scope issued by NICE. We have added text to sections 6.2.1.1 and 6.5.

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

Stakeholder	Comment no.	Page no.	Section no.	Comment	EAG Response
				include GP acceptance, engagement and training and whether follow up of non-responders occurs. To note, data for only OC-Sensor and HM-JACKarc is included so assumptions must be made for the other tests regarding patient acceptability of the sampling devices as a factor that impacts uptake.	
Mast Group Ltd	8	106	Table 19	Table 19: Summary sensitivity and specificity at selected thresholds For those tests with only 1 evidence base we suggest that [b] is expanded to include the study type or population base and number of samples as there is only one data point to work with so how can direct comparison be made with other tests that are supported by more studies.	Thank you, we agree and have implemented this change. We have also added text to state "NB for these tests, the error is the 95% confidence interval for a single study, rather than the 95% CrI of the summary estimate from the meta-analysis model, and these should not be directly compared"
Mast Group Ltd	9	134	4.3.14.1	Test failure rates "There was no strong evidence that rates differ according to test brand." It should be noted in the document that this statement is only applicable to the brands covered in the supporting studies. This sentence should be updated to this effect.	Thank you, this change has been made.
Mast Group Ltd	10	General	General	Can it be explained why some assays have been included without any supporting data? We assume the NICE guidelines will reflect this when published.	The EAG sought data for all studies included in the scope, which was set by NICE. Where we have not been able to identify data, this is clear in our report.

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

Section B Economic model - Comments

Stakeholder	Comment	Description of problem	Description of proposed amendment	Result of amended model or expected impact on the result (if applicable)	EAG response
Mast Group Ltd	1	Tabulated results for each test include a lot of data that could be interpretated to define the thresholds applicable to each test type. Negative predictive value (NPV) and Positive predictive value (PPV) would add some further context to the data and is am important part of decision making for the thresholds.	NPV and PPV for CRC/AAs/LBD should be included in the data tables for ease of interpretation as part of the cost effectiveness	A better picture of clinical implications of the various thresholds within the model will be achieved as part of efforts to increase CRC /AA/ IBD detection.	PPV and NPV have been calculated for all tests at selected, available thresholds for CRC, AA and IBD. These are included in Table 82 in Appendix 5. The prevalence used was as used in the model for the whole cohort, i.e. based on the meta-analysed values from the clinical review for studies of type 2&3 weighted according to D'Souza 2020a.
Mast Group Ltd	2	Within Table 48: Test costs assumed in EAG analysis. For OC-Sensor the accompanying comment below does not make sense and could be clarified better to prevent misinterpretation. '£4.53 Total cost based on the cost per test and sampling test provided by the manufacturer.'	Comment amended to: Total cost included reagent rental of the analyser and the cost per test is indicative. Costs vary depending on testing volumes and methodology employed by the testing laboratory.	No impact on the model.	Amendments to the text in table 48 and the main text of the report were made to reflect the company's suggestion.

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

Stakeholder	Comment	Description of problem	Description of proposed amendment	Result of amended model or expected impact on the result (if applicable)	EAG response
Mast Group Ltd	3	"5.3.9 Deterministic scenario analyses. The EAG has run eleven scenario analyses. For illustrative purposes, the sensitivity analyses have all been conducted on the comparison between HM JACKarc using one threshold of 10µg/g (Intervention 1), in comparison to current recommendations (Intervention 3), using the lower intensity option for safety netting. The summary of results is presented in Table 62, whilst full tables are presented in Table 63 to Table 73 " Why have the scenarios been modelled against just one brand/test – the HMJACKarc? This gives the perception of bias and potentially a commercial advantage to this assay.	Deterministic scenario analyses should be performed for all applicable assays. If this is not feasible some commentary about why that assay was picked for the scenarios and also whether any significant difference between this and the other assays was to be expected for each scenario. It should be clear that there is no bias being shown to one assay over another.	No impact on model unless differences in the scenario outcomes between brands will need to be noted.	We decided to run the scenario analyses on one example test for two reasons. The first is practical: each analysis produces lengthy and multiple tables of results and would increase the size of the report by a large amount. The second is that the evidence base does not allow us to robustly differentiate between the different tests and thresholds. HM-JACKarc was chosen as an example to run the scenario analyses on, and no bias or commercial advantage was intended, or we believe introduced, by this choice. Running the analyses on additional tests would have extremely low value and increase the size and complexity of the report unnecessarily. The results of the scenario analyses are generalisable across tests.

Addendum to the Technology Assessment Report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence – Diagnostics Assessment Report Guide

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

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Sue Harnan, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Aline Navega Biz, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK
Correspondence Author	Sue Harnan, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK
Date completed	19/06/2023

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135637.

1. Introduction

This addendum provides the updated results for the EAG economic analyses presented in the EAG Diagnostics Assessment Report, following the stakeholders' comments where the price for HM-JACKarc test was updated to £4.10 (included in the comments from the manufacturer). All the analyses in this addendum also includes an additional change to the method used to estimate unit costs for FIT in Intervention 3 (use of the weighted mean instead of the minimum value). Please note that due to this change, scenario analysis number 10 now applies the minimum value for FIT in Intervention 3 as the alternative method.

In addition, the EAG has revised the approach adopted for scenario analysis 4 (inclusion of a QALY loss equivalent of one day of full health for each month of diagnostic delay). This analysis includes the correction of programming error which was identified after the final report was submitted, and also includes patients without any significant underlying bowel disease. In the revised approach, the mean times to diagnosis for each underlying disease (CRC, IBD or AA) generated by the model are applied to each intervention, and for patients without any significant underlying bowel disease the model assumes that they would incur a loss of health due to the uncertainty of their (lack of) diagnosis for a mean time of 3 months.

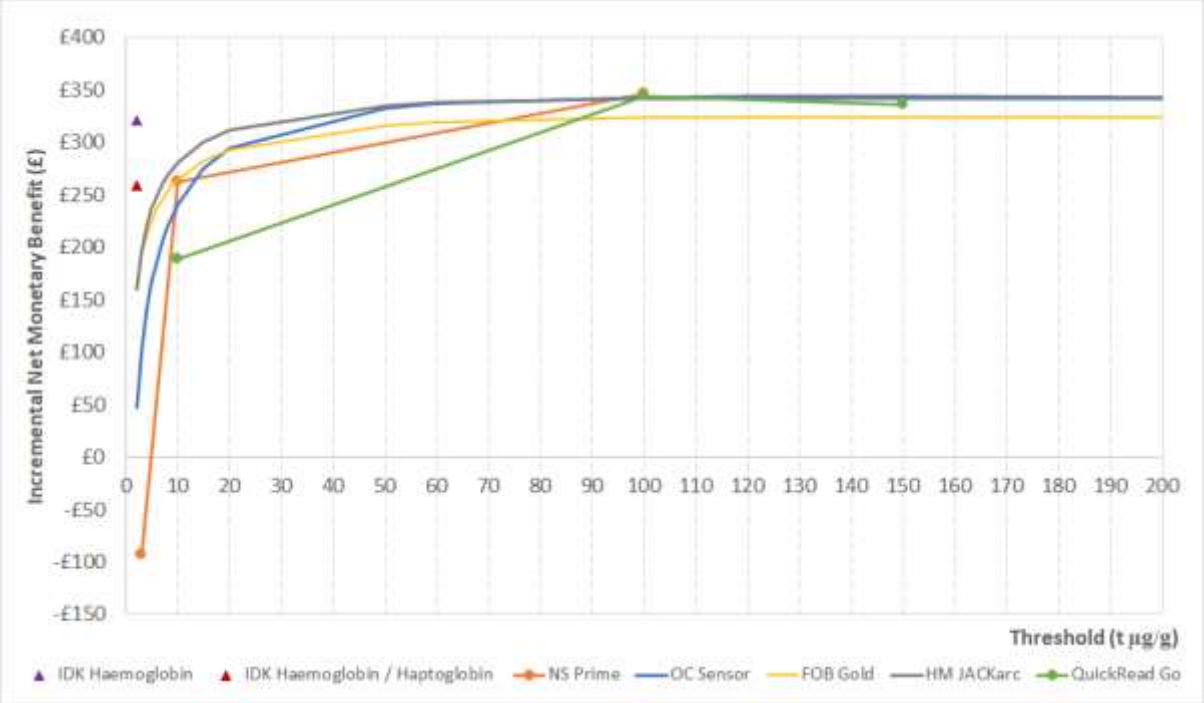
The EAG has also performed two additional analyses (additional scenarios 1 and 2), where the prevalence for CRC, AAs and IBDs were reduced by 50% and increased by 50%.

2. Updated results for the EAG cost-effectiveness results

The EAG updated the results included in the EAG report, which include high or low safety netting intensity and assume a willingness-to-pay threshold of £20,000 or £30,000 per QALY gained. The figures which present incremental net monetary benefit (iNMB) for the low safety netting approach and a threshold of £20,000 per QALY gained are presented in Section 2.1, whilst the tables with results for each test assuming a low safety netting approach are presented in Section 2.2. The results for the scenario analyses run by the EAG (Section 5.3.9 of the EAG report) are shown in Section 3, whilst the results for the analyses originally in Appendix 14 of the EAG report are presented in Section 4 (this includes figures of iNMBs using a threshold of £20,000 plus high safety netting and £30,000 per QALY for both safety netting approaches, and results tables for each test with a high safety netting approach). The results of the two additional scenarios run by the EAG are presented in Section 5.

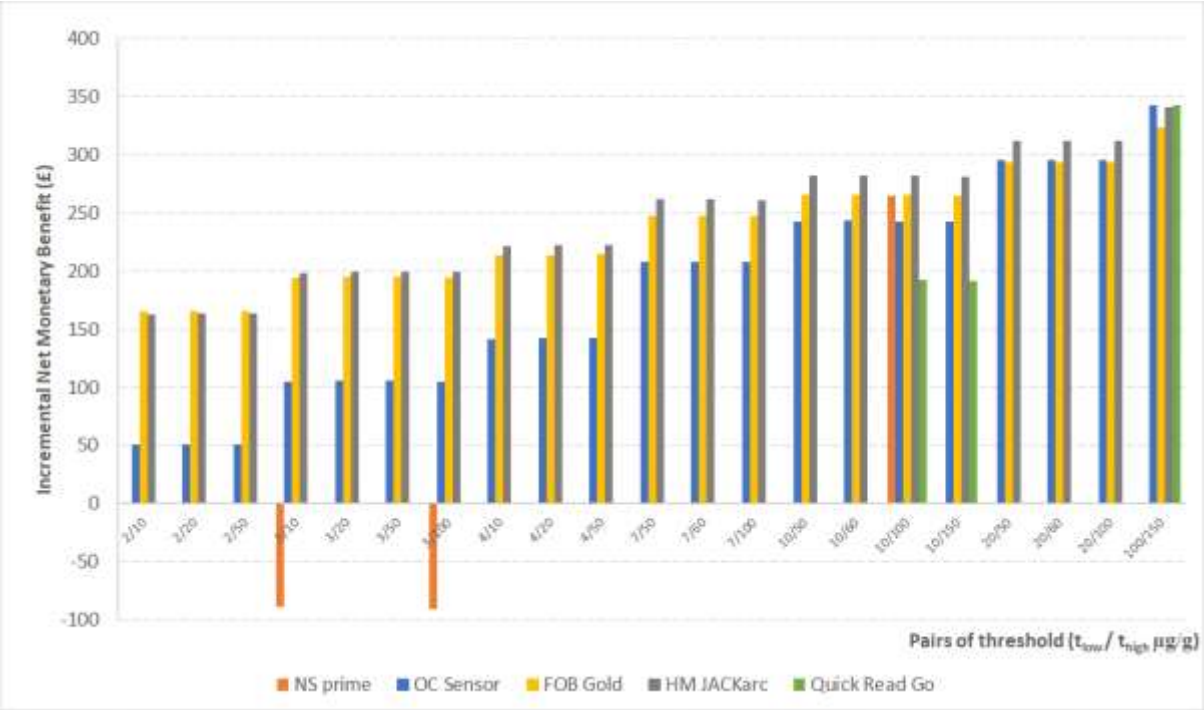
2.1. The iNMBs of the seven tests using one threshold and two thresholds

Figure 1: NMB for Intervention 1 assuming a threshold of £20,000 per QALY gained and low intensity safety netting (Figure 18 of the EAG report)



t – threshold

Figure 2: NMB for Intervention 2 assuming a threshold of £20,000 per QALY gained and low intensity safety netting (Figure 19 of the EAG report)



t_{low} – lower threshold; t_{high} – higher threshold

2.2. Tabulated results for each test (Table 52 to Table 61 of the EAG report)

Table 1: Tabulated results for FOB Gold using one threshold (Table 52 of the EAG report)

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.167	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.894	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2954	2899	2883	2859	2836	2795	2751	2743	2723	3143
ICER (pairwise, vs Intervention 3)□ (£)	146,335	155,625	154,706	150,441	142,991	126,946	103,066	98,530	86,832	-
NMB λ=20,000 (vs Int 3) (£)	163	212	226	246	264	293	316	319	323	-
NMB λ=30,000 (vs Int 3) (£)	150	196	210	227	243	266	278	278	275	-
Number of 2WW referrals (total)	0.385	0.325	0.307	0.282	0.256	0.213	0.168	0.160	0.141	0.644
Number of 18WW referrals (total)	0.065	0.071	0.073	0.076	0.078	0.083	0.088	0.088	0.090	0.037
Number of Repeat FITs (total)	0.065	0.071	0.073	0.076	0.078	0.083	0.088	0.088	0.090	0.037
Number of Watch and Wait (total) (total)	0.485	0.533	0.547	0.567	0.587	0.621	0.657	0.663	0.678	0.281
Number of COLs (total)	0.413	0.364	0.349	0.328	0.307	0.272	0.235	0.229	0.212	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.9%	41.8%	44.2%	47.6%	50.9%	56.5%	62.5%	63.5%	66.1%	-
Reduction in number of referrals (2WW only)	40.2%	49.5%	52.3%	56.2%	60.2%	66.9%	73.9%	75.1%	78.1%	-
Increase in number of referrals (18WW only)□□	72.6%	89.5%	94.5%	101.7%	108.8%	120.9%	133.6%	135.8%	141.3%	-
Reduction in number of COLs	33.8%	41.7%	44.0%	47.4%	50.7%	56.3%	62.3%	63.3%	65.9%	-
Mean time to diagnosis - CRC	2.204	2.339	2.400	2.513	2.668	3.087	3.950	4.169	4.865	1.384
Mean time to diagnosis - AAs	4.453	5.133	5.398	5.833	6.355	7.224	8.346	8.563	9.117	1.956
Mean time to diagnosis - IBD	2.944	3.361	3.521	3.788	4.106	4.814	5.837	6.044	6.607	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 2: Tabulated results for FOB Gold using two thresholds (Table 53 of the EAG report)

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.167	14.167	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.165	14.165	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.892	10.892	10.895
Costs (£)	2951	2949	2948	2897	2896	2894	2855	2854	2832	2831	2793	2792	3143
ICER (pairwise, vs Intervention 3)□ (£)	142,582	139,308	133,327	153,349	150,653	145,561	143,702	139,882	138,255	135,114	124,916	122,824	-
NMB λ=20,000 (vs Int 3) (£)	165	166	166	213	214	214	247	247	265	265	294	294	-
NMB λ=30,000 (vs Int 3) (£)	152	152	151	197	198	197	227	227	243	242	266	265	-
Number of 2WW referrals (total)	0.365	0.358	0.351	0.314	0.308	0.300	0.264	0.259	0.242	0.238	0.206	0.202	0.644
Number of 18WW referrals (total)	0.078	0.083	0.088	0.078	0.083	0.088	0.088	0.090	0.088	0.090	0.088	0.090	0.037
Number of Repeat FITs (total)	0.071	0.074	0.076	0.075	0.077	0.079	0.082	0.083	0.083	0.084	0.085	0.087	0.037
Number of Watch and Wait (total) (total)	0.485	0.485	0.485	0.533	0.533	0.533	0.567	0.567	0.587	0.587	0.621	0.621	0.281
Number of COLs (total)	0.406	0.404	0.402	0.360	0.358	0.356	0.323	0.321	0.303	0.302	0.270	0.269	0.623
Reduction in number of referrals (total - 2WW + 18WW)	34.9%	35.3%	35.6%	42.4%	42.7%	43.1%	48.4%	48.6%	51.6%	51.8%	56.9%	57.1%	-
Reduction in number of referrals (2WW only)	43.3%	44.4%	45.5%	51.2%	52.2%	53.3%	59.0%	59.7%	62.4%	63.0%	68.0%	68.6%	-
Increase in number of referrals (18WW only)	108.8%	120.9%	133.6%	108.8%	120.9%	133.6%	133.6%	141.3%	133.6%	141.3%	133.6%	141.3%	-
Increase in number of repeat FITs	90.7%	96.7%	103.1%	99.2%	105.2%	111.6%	117.7%	121.5%	121.2%	125.1%	127.3%	131.1%	-
Increase in number of watch and waits	72.6%	72.6%	2.6%	89.5%	89.5%	89.5%	101.7%	101.7%	108.8%	108.8%	120.9%	120.9%	-
Reduction in number of COLs	34.8%	35.1%	35.5%	42.2%	42.5%	42.9%	48.3%	48.5%	51.4%	51.6%	56.7%	56.9%	-
Mean time to diagnosis - CRC	2.236	2.265	2.323	2.360	2.386	2.440	2.599	2.654	2.742	2.795	3.135	3.184	1.384
Mean time to diagnosis - AAs	4.580	4.637	4.711	5.209	5.263	5.332	5.980	6.024	6.467	6.511	7.283	7.324	1.956
Mean time to diagnosis - IBD	3.023	3.071	3.141	3.408	3.453	3.518	3.911	3.958	4.207	4.252	4.871	4.913	2.044

□*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 3: Tabulated results for HM JACKarc using one threshold (Table 54 of the EAG report)

	Int 1: FIT 1 threshold (µg/g)									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2955	2890	2870	2843	2817	2774	2730	2723	2705	3143
ICER (pairwise, vs Intervention 3)□ (£)	135,544	151,594	152,020	149,470	143,267	128,674	106,266	101,969	90,930	-
NMB λ=20,000 (vs Int 3) (£)	160	220	237	259	280	312	335	337	341	-
NMB λ=30,000 (vs Int 3) (£)	146	203	219	239	258	283	296	296	293	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)□□	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 4: Tabulated results for HM JACKarc using two thresholds (Table 55 of the EAG report)

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.165	14.165	14.165	14.165	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.892	10.892	10.892	10.892	10.895
Costs (£)	2951	2950	2948	2887	2886	2884	2839	2838	2814	2813	2772	2771	3143
ICER (pairwise, vs Intervention 3) (£)	132,688	129,933	125,005	149,689	147,335	142,978	143,625	140,433	139,143	136,506	126,895	125,118	-
NMB λ=20,000 (vs Int 3) (£)	163	163	163	222	222	223	261	261	282	282	312	312	-
NMB λ=30,000 (vs Int 3) (£)	148	149	148	204	205	204	240	240	258	258	283	283	-
Number of 2WW referrals (total)	0.363	0.355	0.348	0.302	0.295	0.288	0.246	0.242	0.222	0.219	0.184	0.180	0.644
Number of 18WW referrals (total)	0.080	0.085	0.090	0.080	0.085	0.090	0.090	0.092	0.090	0.092	0.090	0.092	0.037
Number of Repeat FITs (total)	0.072	0.075	0.077	0.076	0.079	0.081	0.084	0.085	0.085	0.086	0.087	0.089	0.037
Number of Watch and Wait (total) (total)	0.485	0.485	0.485	0.541	0.541	0.541	0.581	0.581	0.603	0.603	0.638	0.638	0.281
Number of COLs (total)	0.406	0.404	0.402	0.351	0.349	0.347	0.308	0.307	0.287	0.286	0.252	0.251	0.623
Reduction in number of referrals (total - 2WW + 18WW)	35.0%	35.3%	35.7%	43.9%	44.2%	44.6%	50.7%	50.9%	54.2%	54.4%	59.8%	60.0%	-
Reduction in number of referrals (2WW only)	43.7%	44.8%	45.9%	53.1%	54.2%	55.3%	61.8%	62.4%	65.5%	66.1%	71.4%	72.0%	-
Increase in number of referrals (18WW only)	114.4%	127.1%	139.6%	114.4%	127.1%	139.6%	139.6%	146.5%	139.6%	146.5%	139.6%	146.5%	-
Increase in number of repeat FITs	93.4%	99.7%	106.0%	103.4%	109.8%	116.1%	123.1%	126.5%	127.0%	130.5%	133.4%	136.8%	-
Increase in number of watch and waits	72.4%	72.4%	72.4%	92.5%	92.5%	92.5%	106.5%	106.5%	114.4%	114.4%	127.1%	127.1%	-
Reduction in number of COLs	34.8%	35.2%	35.5%	43.7%	44.0%	44.4%	50.5%	50.7%	54.0%	54.2%	59.6%	59.8%	-
Mean time to diagnosis - CRC	2.342	2.371	2.426	2.482	2.508	2.558	2.732	2.779	2.883	2.928	3.279	3.321	1.384
Mean time to diagnosis - AAs	4.579	4.637	4.711	5.200	5.253	5.321	5.963	6.007	6.449	6.491	7.264	7.303	1.956
Mean time to diagnosis - IBD	3.023	3.071	3.140	3.400	3.444	3.508	3.896	3.941	4.189	4.233	4.852	4.893	2.044

□*Southwest quadrant ICER.*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 5: Tabulated results for IDK Haemoglobin using one threshold (Table 56 of the EAG report)

	Int 1: FIT 1 threshold	Int 3: DG30& NG12
t (µg/g)	**2**	**10**
LYs	14.165	14.168
QALYs	10.893	10.895
Costs (£)	2783	3143
ICER (pairwise, vs Intervention 3)□ (£)	182,663	-
NMB λ=20,000 (vs Int 3) (£)	320	-
NMB λ=30,000 (vs Int 3) (£)	301	-
Number of 2WW referrals (total)	0.201	0.644
Number of 18WW referrals (total)	0.084	0.037
Number of Repeat FITs (total)	0.084	0.037
Number of Watch and Wait (total) (total)	0.631	0.281
Number of COLs (total)	0.263	0.623
Reduction in number of referrals (total - 2WW + 18WW)	58.1%	-
Reduction in number of referrals (2WW only)	68.8%	-
Increase in number of referrals (18WW only) □□	124.4%	-
Reduction in number of COLs	57.9%	-
Mean time to diagnosis - CRC	3.027	1.384
Mean time to diagnosis - AAs	4.328	1.956
Mean time to diagnosis - IBD	2.814	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 6: Tabulated results for IDK Haemoglobin using one threshold (Table 57 of the EAG report)

	Int 1: FIT 1threshold	Int 3: DG30& NG12
t (µg/g)	**2**	**10**
LYs	14.164	14.168
QALYs	10.892	10.895
Costs (£)	2836	3143
ICER (pairwise, vs Intervention 3)□ (£)	127,325	-
NMB λ=20,000 (vs Int 3) (£)	258	-
NMB λ=30,000 (vs Int 3) (£)	234	-
Number of 2WW referrals (total)	0.258	0.644
Number of 18WW referrals (total)	0.078	0.037
Number of Repeat FITs (total)	0.078	0.037
Number of Watch and Wait (total) (total)	0.585	0.281
Number of COLs (total)	0.309	0.623
Reduction in number of referrals (total - 2WW + 18WW)	0.506	-
Reduction in number of referrals (2WW only)	59.9%	-
Increase in number of referrals (18WW only) □□	108.3%	-
Reduction in number of COLs	50.4%	-
Mean time to diagnosis - CRC	3.477	1.384
Mean time to diagnosis - AAs	4.367	1.956
Mean time to diagnosis - IBD	2.855	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 7: Tabulated results for NS Prime (Table 58 of the EAG report)

	Int 1: FIT 1 threshold			Int 2: FIT 2 thresholds			Int 3: DG30&NG12 (FIT T=10)
Threshold - t or t_{low}/t_{high} (µg/g)	**3**	**10**	**100**	**3/10**	**3/100**	**10/100**	**10**
LYs	14.164	14.162	14.160	14.164	14.164	14.162	14.168
QALYs	10.892	10.891	10.889	10.892	10.892	10.891	10.895
Costs (£)	3183	2804	2684	3177	3175	2800	3143
ICER (pairwise, vs Intervention 3) (£)	Dominated	88,298□	81,248□	Dominated	Dominated	87,199□	-
NMB λ=20,000 (vs Int 3) (£)	-93	262	346	-89	-91	265	-
NMB λ=30,000 (vs Int 3) (£)	-119	224	289	-117	-120	225	-
Number of 2WW referrals (total)	0.645	0.226	0.101	0.579	0.559	0.206	0.644
Number of 18WW referrals (total)	0.037	0.081	0.095	0.081	0.095	0.095	0.037
Number of Repeat FITs (total)	0.037	0.081	0.095	0.059	0.066	0.088	0.037
Number of Watch and Wait (total) (total)	0.280	0.611	0.710	0.280	0.280	0.611	0.281
Number of COLs (total)	0.624	0.282	0.180	0.604	0.598	0.276	0.623
Reduction in number of referrals (total - 2WW + 18WW)	-0.2%	54.9%	71.3%	3.1%	4.0%	55.9%	-
Reduction in number of referrals (2WW)	-0.2%	64.9%	84.3%	10.1%	13.1%	68.0%	-
Increase in number of referrals (18WW)	0.4%	-117.4%	-152.5%	117.4%	152.5%	152.5%	-
Increase in number of repeat FITs	0.4%	-117.4%	-152.5%	58.5%	76.0%	135.0%	-
Increase in number of watch and waits	0.4%	-117.4%	-152.5%	-0.4%	-0.4%	117.4%	-
Reduction in number of COLs	-0.1%	54.7%	71.1%	3.2%	4.1%	55.7%	-
Mean time to diagnosis - CRC	3.451	4.513	5.762	3.550	3.660	4.582	1.384
Mean time to diagnosis - AAs	5.035	6.331	9.081	5.154	5.391	6.480	1.956
Mean time to diagnosis - IBD	3.385	4.082	6.572	3.454	3.671	4.222	2.044

□*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 8: Tabulated results for OC-Sensor using one threshold (Table 59 of the EAG report)

	Int 1: FIT 1 threshold (µg/g)									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.161	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	3066	2970	2940	2898	2857	2791	2731	2722	2701	3143
ICER (pairwise, vs Intervention 3)□ (£)	51,910	100,401	110,655	120,778	124,939	121,880	104,222	100,142	89,261	-
NMB λ=20,000 (vs Int 3) (£)	47	138	166	205	240	294	333	337	343	-
NMB λ=30,000 (vs Int 3) (£)	32	121	148	184	218	265	293	295	293	-
Number of 2WW referrals (total)	0.510	0.402	0.369	0.322	0.278	0.209	0.147	0.138	0.117	0.644
Number of 18WW referrals (total)	0.052	0.063	0.066	0.071	0.076	0.083	0.090	0.091	0.093	0.037
Number of Repeat FITs (total)	0.052	0.063	0.066	0.071	0.076	0.083	0.090	0.091	0.093	0.037
Number of Watch and Wait (total) (total)	0.387	0.472	0.498	0.535	0.570	0.625	0.673	0.680	0.697	0.281
Number of COLs (total)	0.514	0.426	0.399	0.361	0.325	0.269	0.218	0.211	0.193	0.623
Reduction in number of referrals (total - 2WW + 18WW)	17.6%	31.8%	36.1%	42.2%	48.0%	57.1%	65.2%	66.4%	69.1%	-
Reduction in number of referrals (2WW only)	20.8%	37.6%	42.8%	50.0%	56.8%	67.6%	77.1%	78.5%	81.8%	-
Increase in number of referrals (18WW only)□□	37.7%	68.1%	77.3%	90.3%	102.7%	122.2%	139.5%	142.0%	147.9%	-
Reduction in number of COLs	17.6%	31.7%	36.0%	42.1%	47.9%	56.9%	65.0%	66.2%	69.0%	-
Mean time to diagnosis - CRC	2.347	2.475	2.536	2.653	2.812	3.247	4.118	4.335	5.004	1.384
Mean time to diagnosis - AAs	4.537	5.187	5.442	5.863	6.372	7.220	8.328	8.544	9.096	1.956
Mean time to diagnosis - IBD	3.031	3.416	3.565	3.818	4.122	4.811	5.820	6.025	6.586	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 9: Tabulated results for OC-Sensor using two thresholds (Table 60 of the EAG report)

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.166	14.165	14.165	14.165	14.165	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.892	10.892	10.892	10.892	10.895
Costs (£)	3061	3060	3059	2966	2965	2963	2892	2891	2852	2851	2789	2788	3143
ICER (pairwise, vs Intervention 3)[□] (£)	52,857	52,238	50,479	100,022	98,779	95,831	116,576	113,719	121,652	119,130	120,290	118,498	-
NMB λ=20,000 (vs Int 3) (£)	51	51	51	141	142	142	208	208	243	243	295	295	-
NMB λ=30,000 (vs Int 3) (£)	35	35	34	123	124	124	186	186	219	218	266	265	-
Number of 2WW referrals (total)	0.473	0.462	0.453	0.382	0.371	0.361	0.295	0.290	0.257	0.253	0.199	0.194	0.644
Number of 18WW referrals (total)	0.076	0.083	0.090	0.076	0.083	0.090	0.090	0.093	0.090	0.093	0.090	0.093	0.037
Number of Repeat FITs (total)	0.064	0.067	0.071	0.069	0.073	0.076	0.081	0.082	0.083	0.084	0.087	0.088	0.037
Number of Watch and Wait (total) (total)	0.387	0.387	0.387	0.472	0.472	0.472	0.535	0.535	0.570	0.570	0.625	0.625	0.281
Number of COLs (total)	0.503	0.499	0.496	0.420	0.416	0.413	0.353	0.351	0.319	0.317	0.266	0.264	0.623
Reduction in number of referrals (total - 2WW + 18WW)	19.4%	19.9%	20.4%	32.8%	33.3%	33.8%	43.6%	43.8%	49.0%	49.3%	57.6%	57.8%	-
Reduction in number of referrals (2WW only)	26.5%	28.2%	29.7%	40.7%	42.4%	43.9%	54.2%	55.0%	60.0%	60.7%	69.1%	69.8%	-
Increase in number of referrals (18WW only)	102.7%	122.2%	139.5%	102.7%	122.2%	139.5%	139.5%	147.9%	139.5%	147.9%	139.5%	147.9%	-
Increase in number of repeat FITs	70.2%	79.9%	88.6%	85.4%	95.1%	103.8%	114.9%	119.1%	121.1%	125.3%	130.8%	135.0%	-
Increase in number of watch and waits	37.7%	37.7%	37.7%	68.1%	68.1%	68.1%	90.3%	90.3%	102.7%	102.7%	122.2%	122.2%	-
Reduction in number of COLs	19.4%	19.9%	20.4%	32.7%	33.2%	33.7%	43.4%	43.7%	48.9%	49.1%	57.4%	57.6%	-
Mean time to diagnosis - CRC	2.385	2.419	2.486	2.499	2.529	2.589	2.745	2.800	2.890	2.942	3.294	3.342	1.384
Mean time to diagnosis - AAs	4.678	4.743	4.827	5.268	5.326	5.400	6.015	6.062	6.485	6.530	7.278	7.319	1.956
Mean time to diagnosis - IBD	3.118	3.171	3.249	3.466	3.513	3.583	3.945	3.993	4.224	4.270	4.866	4.908	2.044

[□]*Southwest quadrant ICER. - AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;*

Table 10: Tabulated results for QuikRead go (Table 61 of the EAG report)

	Int 1: FIT 1 threshold			**Int 2: FIT 2 thresholds**			**Int 3: DG30& NG12**
t (µg/g)	**10**	**100**	**150**	**10/100**	**10/150**	**100/150**	**10**
LYs	14.166	14.163	14.160	14.166	14.166	14.162	14.168
QALYs	10.893	10.890	10.889	10.893	10.892	10.890	10.895
Costs (£)	2913	2710	2692	2906	2905	2710	3143
ICER (pairwise, vs Intervention 3)□ (£)	110,463	97,506	78,336	105,942	102,381	96,304	-
NMB λ=20,000 (vs Int 3) (£)	188	344	336	192	191	343	-
NMB λ=30,000 (vs Int 3) (£)	168	299	278	170	168	298	-
Number of 2WW referrals (total)	0.338	0.126	0.110	0.305	0.302	0.123	0.644
Number of 18WW referrals (total)	0.070	0.092	0.094	0.092	0.094	0.094	0.037
Number of Repeat FITs (total)	0.070	0.092	0.094	0.081	0.082	0.093	0.037
Number of Watch and Wait (total) (total)	0.523	0.690	0.703	0.523	0.523	0.690	0.281
Number of COLs (total)	0.374	0.200	0.187	0.364	0.363	0.200	0.623
Reduction in number of referrals (total - 2WW + 18WW)	40.2%	68.0%	70.1%	41.8%	41.9%	68.2%	-
Reduction in number of referrals (2WW only)	47.5%	80.5%	82.9%	52.7%	53.1%	80.8%	-
Increase in number of referrals (18WW only) □□	85.9%	145.5%	149.9%	-145.5%	-149.9%	-149.9%	-
Reduction in number of COLs	40.0%	67.9%	70.0%	41.7%	41.8%	68.0%	-
Mean time to diagnosis - CRC	2.557	4.438	5.770	2.677	2.762	4.502	1.384
Mean time to diagnosis - AAs	6.418	9.104	9.484	6.586	6.610	9.121	1.956
Mean time to diagnosis - IBD	4.167	6.593	7.027	4.324	4.352	6.615	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

3. Updated results for the EAG's deterministic scenario analyses (Table 62 to Table 73 of the EAG report)

Table 11: Deterministic sensitivity analyses results for HM JACKarc using one threshold (10 µg/g) (Table 62 of the EAG report)

	Intervention 1 (FIT using threshold of 10) versus Intervention 3 (DG30/NG12)			
Scenario	**Inc. QALYs**	**Inc. costs**	**ICER□**	**iNMB (20k)**
Base case (deterministic)	-0.0023	-326	143,267	280
Scenario 1: shorter time to diagnosis (best-case)	-0.0012	-325	275,179	302
Scenario 2: longer time to diagnosis (worst-case)	-0.0041	-326	78,704	244
Scenario 3: QALY loss due to receiving a colonoscopy	-0.0015	-326	212,440	295
Scenario 4: QALY loss for each month of diagnostic delay	-0.0027	-326	120,071	271
Scenario 5: DUAL FIT	-0.0015	-229	151,012	198
Scenario 6: removing IBD and AAs from the model	-0.0012	-364	305,119	340
Scenario 7: Using alternative source for FIT return rate from Moss *et al*. (2017)[124]	-0.0043	-321	83,885	275
Scenario 8: Use of accuracy data for DG30 low-risk group (Intervention 3) from EAG's clinical review analysis for this group	-0.0023	-262	128,481	242
Scenario 9: Increased resource use of GP appointments for patients with NSBP following watch and wait or Repeat FIT	-0.0023	-288	137,832	268
Scenario 10: Alternative method to estimate unit costs for FIT in Intervention 3 (weighted mean)	-0.0023	-325	143,146	280
Scenario 11: FIT has perfect accuracy (sensitivity and specificity =1.0) and return rate =1.0	0.0007	-440	Dominates	453

□*Southwest quadrant ICER*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; QALY - quality-adjusted life year

Scenario 1: shorter time-to-diagnosis (best-case)

Table 12: Tabulated results for HM JACKarc using one threshold (Table 63 of the EAG report)

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.168	14.168	14.168	14.168	14.168	14.167	14.166	14.166	14.165	14.169
QALYs	10.895	10.894	10.894	10.894	10.894	10.894	10.893	10.893	10.893	10.895
Costs (£)	2956	2891	2871	2844	2818	2775	2732	2725	2708	3143
ICER (pairwise, vs Intervention 3)□ (£)	263,597	296,243	296,375	289,651	275,179	242,651	195,585	186,846	164,750	-
iNMB λ=20,000 (vs Int 3) (£)	173	236	254	278	302	338	369	373	383	-
iNMB λ=30,000 (vs Int 3) (£)	166	227	244	268	290	323	348	351	356	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)□□	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	1.519	1.591	1.626	1.692	1.781	2.018	2.481	2.595	2.947	1.044
Mean time to diagnosis - AAs	2.821	3.194	3.342	3.586	3.881	4.374	5.016	5.141	5.460	1.444
Mean time to diagnosis - IBD	1.863	2.087	2.175	2.324	2.503	2.909	3.501	3.622	3.951	1.396

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 2: longer time to diagnosis (worst-case)

Table 13: Tabulated results for HM JACKarc using one threshold (Table 64 of the EAG report)

	Int 1: FIT 1 threshold									**Int 3 : DG30& NG12**
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.163	14.163	14.162	14.162	14.161	14.160	14.157	14.157	14.155	14.166
QALYs	10.891	10.890	10.890	10.890	10.889	10.888	10.886	10.886	10.885	10.893
Costs (£)	2953	2887	2868	2841	2815	2771	2726	2719	2700	3,141
ICER (pairwise, vs Intervention 3)□ (£)	74,022	82,759	83,076	81,863	78,704	71,149	59,292	56,995	51,067	-
iNMB λ=20,000 (vs Int 3) (£)	137	193	207	227	244	266	275	274	268	-
iNMB λ=30,000 (vs Int 3) (£)	112	162	174	190	202	214	205	200	182	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)□□	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	4.170	4.449	4.572	4.795	5.087	5.842	7.281	7.630	8.709	2.483
Mean time to diagnosis - AAs	8.151	9.340	9.804	10.565	11.480	13.002	14.971	15.351	16.325	3.702
Mean time to diagnosis - IBD	5.232	5.970	6.253	6.728	7.293	8.556	10.379	10.748	11.753	3.585

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 3: QALY loss due to receiving a colonoscopy

Table 14: Tabulated results for HM JACKarc using one threshold (Table 65 of the EAG report)

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.892	10.892	10.892	10.892	10.891	10.890	10.890	10.890	10.893
Costs (£)	2955	2890	2870	2843	2817	2774	2730	2723	2705	3,143
ICER (pairwise, vs Intervention 3)□ (£)	204,856	236,303	235,772	227,877	212,440	180,463	138,521	131,197	113,262	-
NMB λ=20,000 (vs Int 3) (£)	169	232	249	273	295	328	353	356	360	-
NMB λ=30,000 (vs Int 3) (£)	160	221	238	260	280	308	323	324	322	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)□□	-72.4%	-92.5%	-98.3%	-106.5%	-114.4%	-127.1%	-139.6%	-141.6%	-146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 4: QALY loss for each month of diagnostic delay

Table 15: Tabulated results for HM JACKarc using one threshold (Table 66 of the EAG report)

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.887	10.886	10.886	10.886	10.886	10.885	10.884	10.883	10.883	10.888
Costs (£)	2955	2890	2870	2843	2817	2774	2730	2723	2705	3143
ICER (pairwise, vs Intervention 3)□ (£)	115,455	127,877	127,946	125,479	120,071	107,858	89,564	86,075	77,129	-
NMB λ=20,000 (vs Int 3) (£)	155	214	230	252	271	300	320	322	324	-
NMB λ=30,000 (vs Int 3) (£)	139	194	209	228	244	266	274	273	267	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)□□	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 5: DUAL FIT

Table 16: Tabulated results for HM JACKarc using one threshold (Table 67 of the EAG report)

	Int 1: FIT 1 threshold	Int 3 : DG30& NG12
t (µg/g)	10	10
LYs	14.167	14.168
QALYs	10.893	10.895
Costs (£)	2914	3143
ICER (pairwise, vs Intervention 3)□ (£)	151,012	-
NMB λ=20,000 (vs Int 3) (£)	198	-
NMB λ=30,000 (vs Int 3) (£)	183	-
Number of 2WW referrals (total)	0.336	0.644
Number of 18WW referrals (total)	0.070	0.037
Number of Repeat FITs (total)	0.070	0.037
Number of Watch and Wait (total) (total)	0.524	0.281
Number of COLs (total)	0.373	0.623
Reduction in number of referrals (total - 2WW + 18WW)	40.4%	-
Reduction in number of referrals (2WW only)	47.7%	-
Increase in number of referrals (18WW only)□□	86.3%	-
Reduction in number of COLs	40.2%	-
Mean time to diagnosis - CRC	2.204	1.384
Mean time to diagnosis - AAs	5.512	1.956
Mean time to diagnosis - IBD	2.397	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 6: removing IBD and AA from the model

Table 17: Tabulated results for HM JACKarc using one threshold (Table 68 of the EAG report)

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.193	14.193	14.192	14.192	14.192	14.192	14.191	14.190	14.189	14.195
QALYs	10.944	10.944	10.944	10.944	10.944	10.944	10.943	10.943	10.942	10.945
Costs (£)	932	861	841	812	784	739	695	688	670	1,148
ICER (pairwise, vs Intervention 3)□ (£)	281,084	323,282	325,635	320,255	305,119	260,284	196,467	185,119	156,997	-
NMB λ=20,000 (vs Int 3) (£)	200	269	288	315	340	377	407	410	417	-
NMB λ=30,000 (vs Int 3) (£)	193	260	279	305	328	362	384	386	386	-
Number of 2WW referrals (total)	0.365	0.293	0.272	0.243	0.216	0.189	0.172	0.130	0.124	0.639
Number of 18WW referrals (total)	0.067	0.074	0.077	0.080	0.083	0.085	0.087	0.092	0.092	0.038
Number of Repeat FITs (total)	0.067	0.074	0.077	0.080	0.083	0.085	0.087	0.092	0.092	0.038
Number of Watch and Wait (total) (total)	0.501	0.558	0.575	0.598	0.619	0.641	0.654	0.686	0.692	0.285
Number of COLs (total)	0.394	0.335	0.318	0.295	0.272	0.250	0.237	0.203	0.198	0.617
Reduction in number of referrals (total - 2WW + 18WW)	36.2%	45.8%	48.5%	52.4%	56.0%	59.5%	61.7%	67.2%	68.1%	-
Reduction in number of referrals (2WW only)	42.8%	54.2%	57.5%	62.0%	66.3%	70.5%	73.1%	79.6%	80.6%	-
Increase in number of referrals (18WW only)□□	75.9%	96.1%	101.8%	109.9%	117.4%	124.9%	129.6%	141.0%	142.8%	-
Reduction in number of COLs	36.1%	45.7%	48.4%	52.3%	55.9%	61.6%	67.1%	68.0%	70.0%	-
Mean time to diagnosis - CRC	2.296	2.447	2.514	2.638	2.801	3.030	3.224	4.033	4.230	1.384
Mean time to diagnosis - AAs	-	-	-	-	-	-	-	-	-	-
Mean time to diagnosis - IBD	-	-	-	-	-	-	-	-	-	-

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 7: Using alternative source for FIT return rate from Moss et al. (2017)[124]

Table 18: Tabulated results for HM JACKarc using one threshold (Table 69 of the EAG report)

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.162	14.162	14.162	14.162	14.161	14.161	14.160	14.159	14.159	14.167
QALYs	10.891	10.891	10.891	10.890	10.890	10.890	10.889	10.889	10.888	10.894
Costs (£)	2858	2809	2795	2775	2755	2723	2691	2686	2673	3,116
ICER (pairwise, vs Intervention 3)□ (£)	70,311	79,334	81,209	83,119	83,885	83,019	77,685	76,270	72,101	-
NMB λ=20,000 (vs Int 3) (£)	184	229	242	259	275	298	315	317	320	-
NMB λ=30,000 (vs Int 3) (£)	148	191	202	218	232	251	261	261	259	-
Number of 2WW referrals (total)	0.296	0.244	0.229	0.207	0.187	0.154	0.121	0.116	0.103	0.622
Number of 18WW referrals (total)	0.074	0.080	0.081	0.083	0.086	0.089	0.093	0.093	0.094	0.040
Number of Repeat FITs (total)	0.074	0.080	0.081	0.083	0.086	0.089	0.093	0.093	0.094	0.040
Number of Watch and Wait (total) (total)	0.555	0.597	0.609	0.626	0.642	0.668	0.694	0.698	0.708	0.299
Number of COLs (total)	0.340	0.297	0.284	0.267	0.250	0.223	0.196	0.192	0.181	0.605
Reduction in number of referrals (total - 2WW + 18WW)	44.0%	51.1%	53.1%	56.0%	58.8%	63.3%	67.7%	68.4%	70.2%	-
Reduction in number of referrals (2WW only)	52.3%	60.8%	63.2%	66.6%	69.9%	75.3%	80.5%	81.4%	83.4%	-
Increase in number of referrals (18WW only)□□	85.9%	99.8%	103.8%	109.4%	114.9%	123.6%	132.3%	133.7%	137.0%	-
Reduction in number of COLs	43.9%	51.0%	53.0%	55.9%	58.7%	63.2%	67.6%	68.3%	70.1%	-
Mean time to diagnosis - CRC	4.683	4.787	4.834	4.922	5.038	5.344	5.932	6.075	6.519	1.760
Mean time to diagnosis - AAs	6.285	6.771	6.961	7.275	7.653	8.284	9.104	9.262	9.668	1.956
Mean time to diagnosis - IBD	5.132	5.424	5.537	5.728	5.957	6.470	7.216	7.367	7.779	2.757

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 8: Use of accuracy data for DG30 low-risk group (Intervention 3) from EAG's clinical review analysis for this group

Table 19: Tabulated results for HM JACKarc using one threshold (Table 70 of the EAG report)

	Int 1: FIT 1 threshold									**Int 3 : DG30& NG12**
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2955	2890	2870	2843	2817	2774	2730	2723	2705	3105
ICER (pairwise, vs Intervention 3)□ (£)	105,436	126,338	128,481	128,488	124,880	114,277	95,823	92,151	82,602	-
NMB λ=20,000 (vs Int 3) (£)	122	182	198	221	242	273	297	299	303	-
NMB λ=30,000 (vs Int 3) (£)	107	165	180	201	219	244	258	258	255	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.604
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.042
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.042
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.312
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.591
Reduction in number of referrals (total - 2WW + 18WW)	30.2%	40.1%	43.0%	47.1%	50.9%	57.2%	63.4%	64.4%	66.8%	-
Reduction in number of referrals (2WW only)	42.8%	54.2%	57.5%	62.0%	66.3%	70.5%	73.1%	79.6%	80.6%	-
Increase in number of referrals (18WW only)□□	75.9%	96.1%	101.8%	109.9%	117.4%	124.9%	129.6%	141.0%	142.8%	-
Reduction in number of COLs	30.1%	40.0%	42.8%	46.9%	50.7%	57.0%	63.2%	64.2%	66.6%	-
Mean time to diagnosis - CRC	2.325	2.475	2.542	2.665	2.826	3.248	4.055	4.251	4.858	1.375
Mean time to diagnosis - AAs	4.472	5.143	5.405	5.836	6.355	7.220	8.341	8.557	9.112	1.956
Mean time to diagnosis - IBD	2.964	3.370	3.527	3.791	4.105	4.810	5.831	6.038	6.602	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 9: Increased resource use of GP appointments for patients with NSBP following watch and wait or Repeat FIT

Table 20: Tabulated results for HM JACKarc using one threshold (Table 71 of the EAG report)

	Int 1: FIT 1 threshold									**Int 3 : DG30& NG12**
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2974	2911	2892	2866	2841	2799	2756	2749	2732	3154
ICER (pairwise, vs Intervention 3)□ (£)	129,860	145,574	146,060	143,710	137,832	123,914	102,444	98,320	87,722	-
NMB λ=20,000 (vs Int 3) (£)	152	210	226	248	268	298	320	322	326	-
NMB λ=30,000 (vs Int 3) (£)	138	193	208	228	245	269	281	281	278	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)□□	-72.4%	-92.5%	-98.3%	-106.5%	-114.4%	-127.1%	-139.6%	-141.6%	-146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 10: Alternative method to estimate unit costs for FIT in Intervention 3 (weighted mean)

Table 21: Tabulated results for HM JACKarc using one threshold (Table 72 of the EAG report)

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.166	14.166	14.166	14.166	14.166	14.165	14.163	14.163	14.162	14.168
QALYs	10.893	10.893	10.893	10.893	10.893	10.892	10.891	10.891	10.890	10.895
Costs (£)	2955	2890	2870	2843	2817	2774	2730	2723	2705	3143
ICER (pairwise, vs Intervention 3)□ (£)	135,345	151,429	151,866	149,333	143,146	128,577	106,195	101,902	90,873	-
NMB λ=20,000 (vs Int 3) (£)	160	220	236	259	280	311	335	337	341	-
NMB λ=30,000 (vs Int 3) (£)	146	203	218	239	257	283	296	296	293	-
Number of 2WW referrals (total)	0.386	0.314	0.294	0.265	0.237	0.191	0.147	0.140	0.122	0.644
Number of 18WW referrals (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Repeat FITs (total)	0.065	0.072	0.074	0.077	0.080	0.085	0.090	0.091	0.092	0.037
Number of Watch and Wait (total) (total)	0.485	0.541	0.558	0.581	0.603	0.638	0.674	0.679	0.693	0.281
Number of COLs (total)	0.413	0.355	0.338	0.314	0.291	0.254	0.218	0.212	0.197	0.623
Reduction in number of referrals (total - 2WW + 18WW)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	-
Reduction in number of referrals (2WW only)	40.0%	51.2%	54.4%	58.9%	63.2%	70.3%	77.2%	78.3%	81.0%	-
Increase in number of referrals (18WW only)□□	-72.4%	-92.5%	-98.3%	-106.5%	-114.4%	-127.1%	-139.6%	-141.6%	-146.5%	-
Reduction in number of COLs	33.7%	43.1%	45.8%	49.6%	53.3%	59.2%	65.1%	66.0%	68.4%	-
Mean time to diagnosis - CRC	2.308	2.460	2.527	2.651	2.813	3.236	4.044	4.241	4.848	1.384
Mean time to diagnosis - AAs	4.453	5.126	5.388	5.820	6.340	7.206	8.328	8.545	9.100	1.956
Mean time to diagnosis - IBD	2.944	3.353	3.511	3.775	4.091	4.797	5.819	6.026	6.590	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Scenario 11: FIT with perfect accuracy and return rate=1.0

Table 22: Tabulated results for HM JACKarc using one threshold (Table 73 of the EAG report)

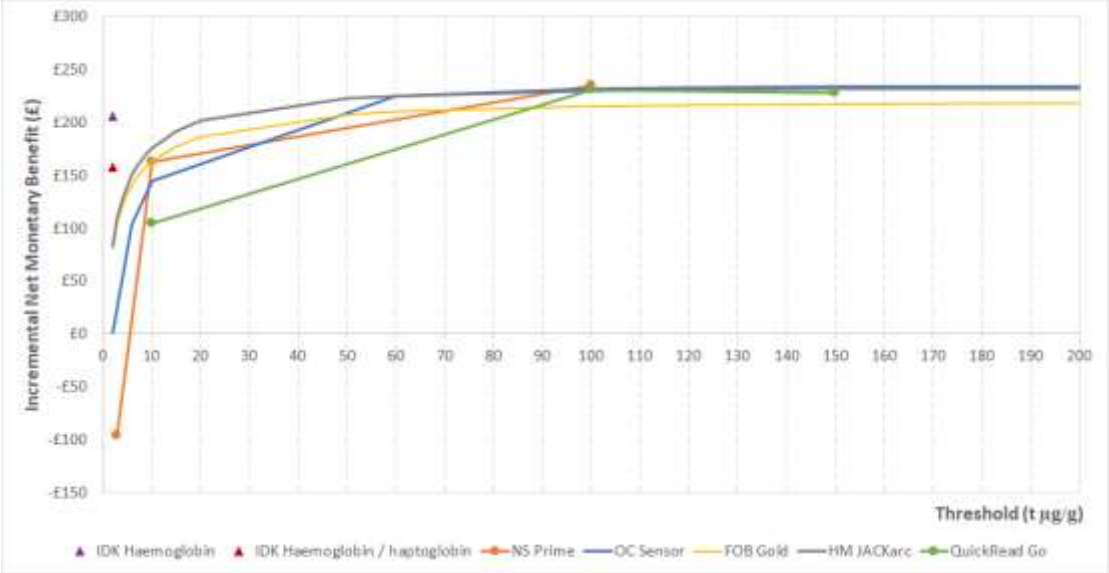
	Int 1: FIT 1 threshold	Int 3 : DG30& NG12
t (µg/g)	-	10
LYs	14.169	14.168
QALYs	10.896	10.895
Costs (£)	2713	3153
ICER (pairwise, vs Intervention 3)□ (£)	Dominates	-
NMB λ=20,000 (vs Int 3) (£)	453	-
NMB λ=30,000 (vs Int 3) (£)	460	-
Number of 2WW referrals (total)	0.124	0.644
Number of 18WW referrals (total)	0.092	0.037
Number of Repeat FITs (total)	0.092	0.037
Number of Watch and Wait (total) (total)	0.691	0.281
Number of COLs (total)	0.201	0.623
Reduction in number of referrals (total - 2WW + 18WW)	68.6%	-
Reduction in number of referrals (2WW only)	81.0%	-
Increase in number of referrals (18WW only)□□	152.1%	-
Reduction in number of COLs	68.1%	-
Mean time to diagnosis - CRC	0.770	1.238
Mean time to diagnosis - AAs	1.668	1.956
Mean time to diagnosis - IBD	0.355	1.766

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

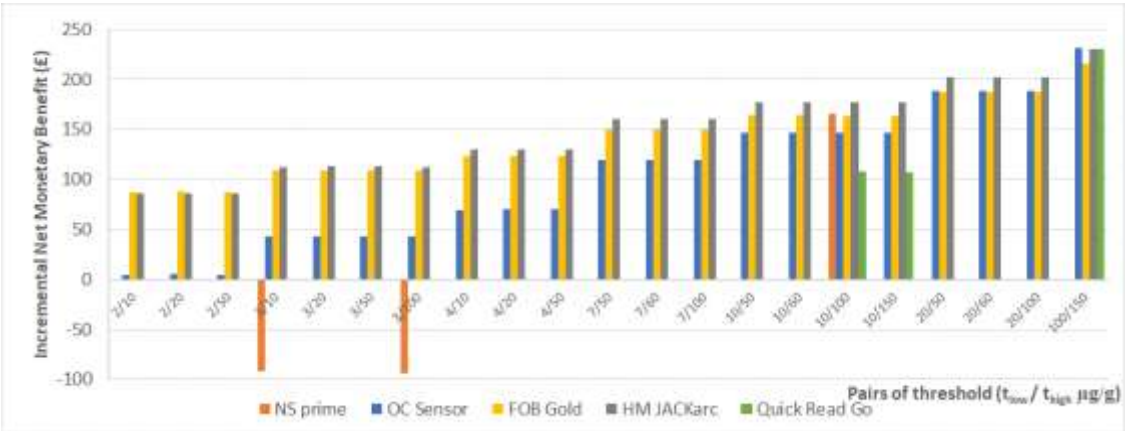
4. Updated results for Additional health economic analyses presented in Appendix 14

Figure 3: ***iNMB for Intervention 1 assuming a threshold of £20,000 per QALY gained and high intensity safety netting*** **(Figure 24 of the EAG report)**



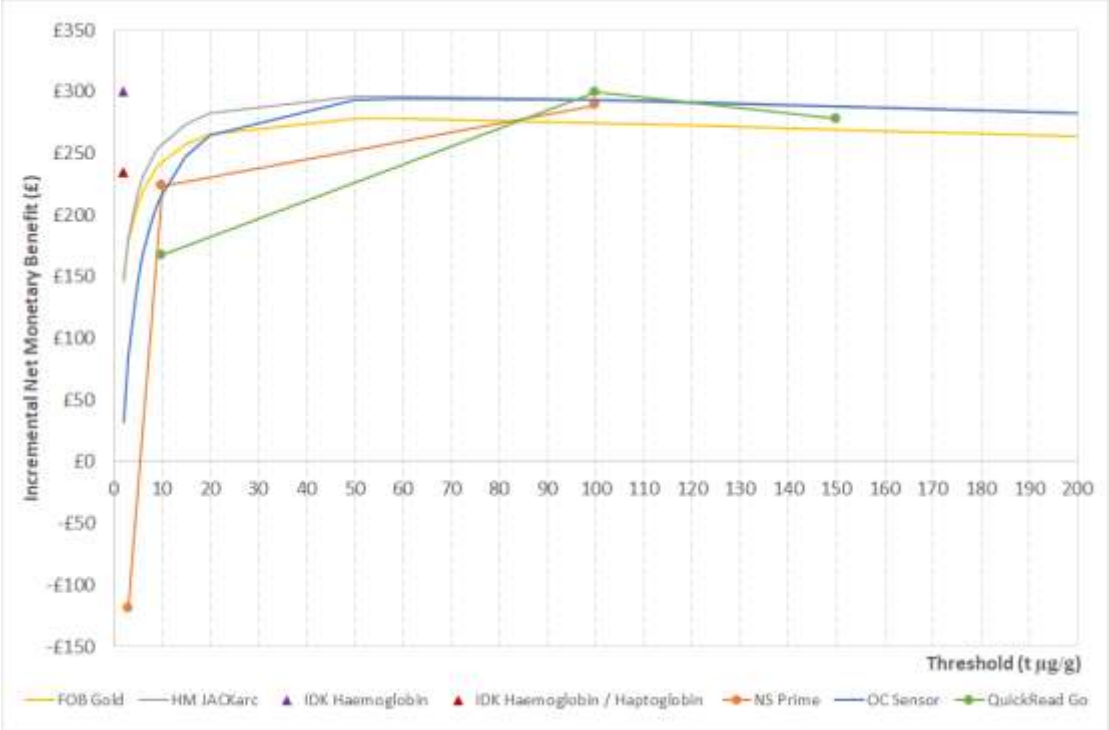
t – threshold

Figure 4: ***iNMB for Intervention 2 assuming a threshold of £20,000 per QALY gained and high intensity safety netting*** **(Figure 25 of the EAG report)**



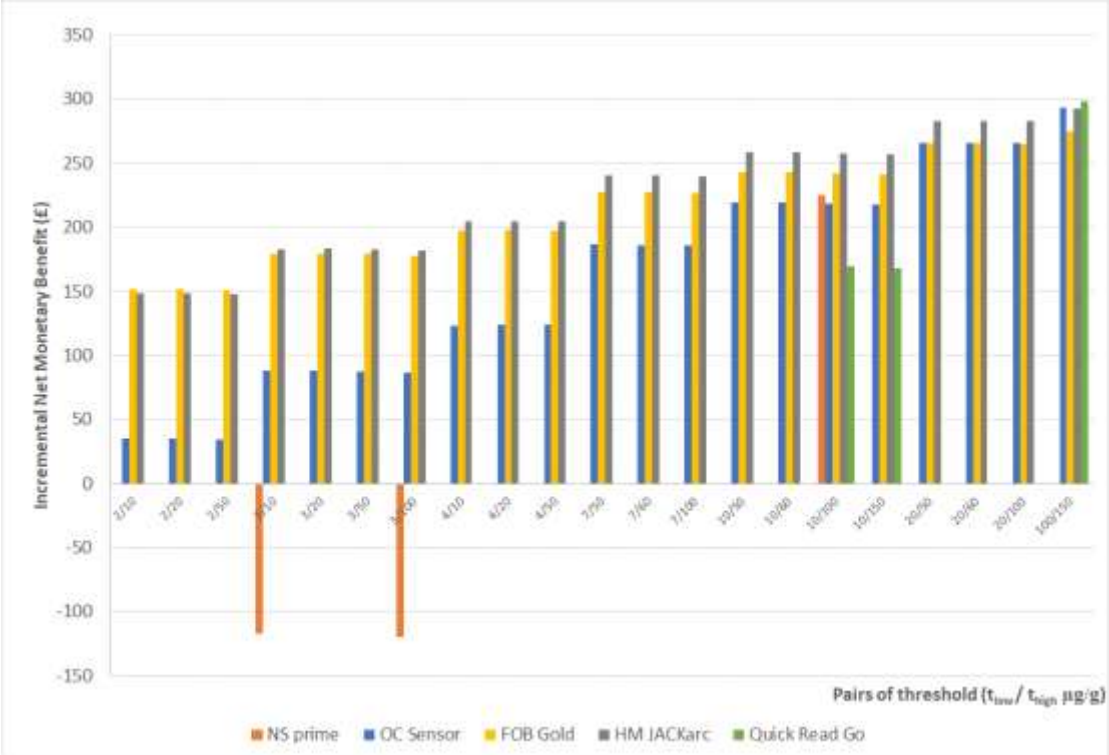
t_{low} – lower threshold; t_{high} – higher threshold

Figure 5: ***iNMB for Intervention 1 assuming a threshold of £30,000 per QALY gained and low intensity safety netting*** **(Figure 26 of the EAG report)**



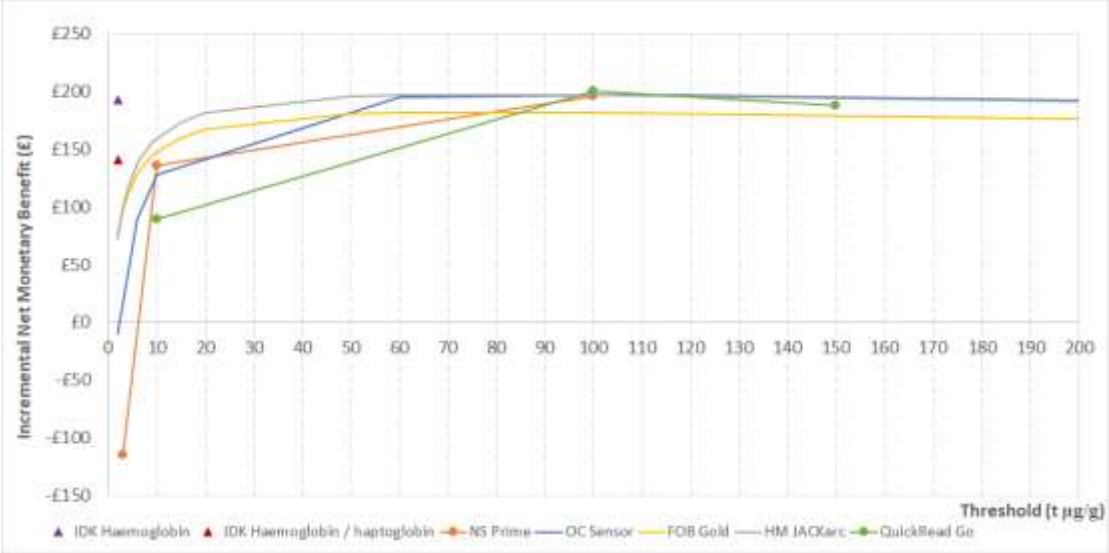
t – threshold

Figure 6: ***iNMB for Intervention 2 assuming a threshold of £30,000 per QALY gained and low intensity safety netting*** **(Figure 27 of the EAG report)**



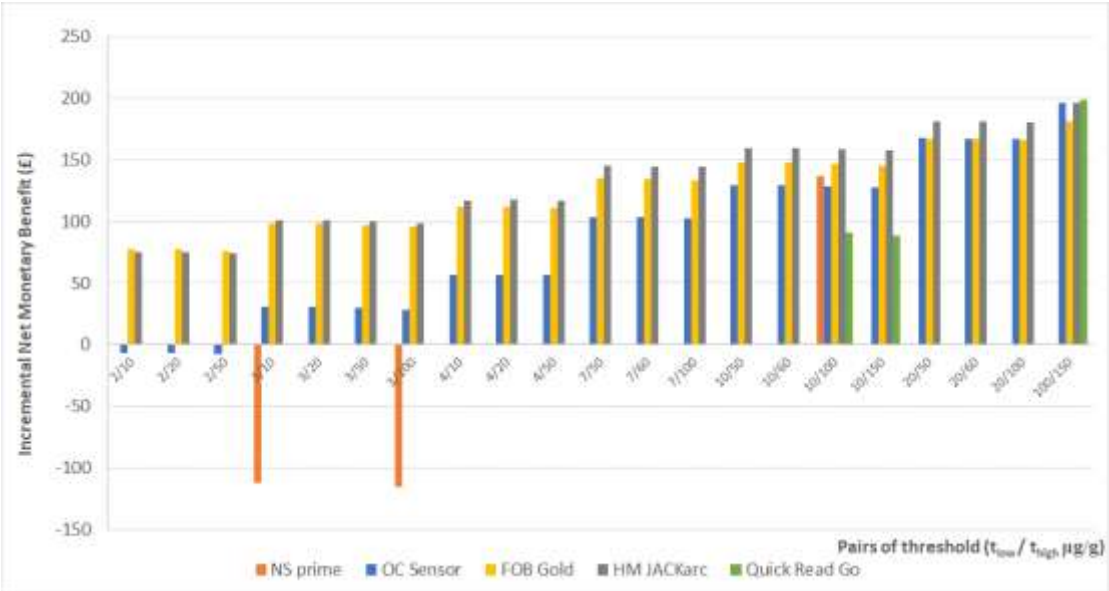
t_{low} *– lower threshold;* t_{high} *– higher threshold*

Figure 7: ***iNMB for Intervention 1 assuming a threshold of £30,000 per QALY gained and high intensity safety netting*** **(Figure 28 of the EAG report)**



t – threshold

Figure 8: ***iNMB for Intervention 2 assuming a threshold of £30,000 per QALY gained and high intensity safety netting*** **(Figure 29 of the EAG report)**



t_{low} *– lower threshold;* t_{high} *– higher threshold*

Tabulated results using high intensity safety netting

Table 23: Tabulated results for FOB Gold using one threshold (Table 92 of the EAG report)

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	**2**	**4**	**5**	**7**	**10**	**20**	**50**	**60**	**100**	**10**
LYs	14.167	14.167	14.167	14.167	14.167	14.166	14.165	14.165	14.164	14.168
QALYs	10.894	10.894	10.894	10.894	10.893	10.893	10.892	10.892	10.891	10.895
Costs (£)	3143	3102	3089	3071	3053	3021	2986	2980	2964	3,247
ICER (pairwise, vs Intervention 3)□ (£)	110,774	128,421	130,284	130,019	126,480	116,248	97,473	93,664	83,583	-
NMB λ=20,000 (vs Int 3) (£)	£85	£122	£133	£149	£163	£187	£207	£210	£215	-
NMB λ=30,000 (vs Int 3) (£)	£76	£111	£121	£135	£148	£167	£181	£182	£182	-
Number of 2WW referrals (total)	0.450	0.560	0.024	0.002	0.001	3.751	0.022	0.249	0.004	0.681
Number of 18WW referrals (total)	0.162	0.178	0.182	0.189	0.196	0.207	0.219	0.221	0.226	0.094
Number of Repeat FITs (total)	0.129	0.142	0.146	0.151	0.157	0.166	0.175	0.177	0.181	0.075
Number of Watch and Wait (total) (total)	0.259	0.284	0.292	0.302	0.313	0.331	0.350	0.354	0.362	0.150
Number of COLs (total)	0.560	0.525	0.515	0.500	0.486	0.461	0.434	0.430	0.419	0.709
Reduction in number of referrals (total - 2WW + 18WW)	21.1%	26.0%	27.4%	29.5%	31.6%	35.1%	38.8%	39.4%	41.0%	-
Reduction in number of referrals (2WW only)	33.9%	41.8%	44.2%	47.6%	50.9%	56.5%	62.5%	63.5%	66.1%	-
Increase in number of referrals (18WW only)□□	72.6%	89.5%	94.5%	101.7%	108.8%	120.9%	133.6%	135.8%	141.3%	-
Reduction in number of COLs	21.0%	25.9%	27.3%	29.4%	31.5%	35.0%	38.7%	39.3%	40.9%	-
Mean time to diagnosis - CRC	1.899	1.991	2.034	2.112	2.219	2.510	3.106	3.257	3.737	1.303
Mean time to diagnosis - AAs	3.751	4.223	4.406	4.706	5.065	5.657	6.418	6.564	6.937	1.956
Mean time to diagnosis - IBD	2.317	2.606	2.717	2.902	3.121	3.609	4.312	4.453	4.838	1.684

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 24: Tabulated results for FOB Gold using two thresholds (Table 93 of the EAG report)

	Int 2: FIT 2 thresholds												Int 3[§]
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.167	14.167	14.167	14.167	14.167	14.167	14.167	14.167	14.166	14.166	14.166	14.166	14.168
QALYs	10.894	10.894	10.894	10.894	10.894	10.894	10.893	10.893	10.893	10.893	10.893	10.893	10.895
Costs (£)	3139	3138	3137	3100	3098	3097	3067	3066	3050	3049	3019	3018	3247
ICER (pairwise, vs Intervention 3)[□] (£)	107,115	103,684	97,476	125,585	122,242	116,064	120,791	115,810	119,594	115,235	112,992	109,766	-
NMB λ=20,000 (vs Int 3) (£)	87	88	87	123	124	124	150	149	164	164	187	187	-
NMB λ=30,000 (vs Int 3) (£)	77	77	76	112	112	111	135	134	148	147	167	166	-
Number of 2WW referrals (total)	0.430	0.423	0.416	0.385	0.379	0.371	0.339	0.335	0.321	0.316	0.289	0.285	0.681
Number of 18WW referrals (total)	0.175	0.180	0.185	0.185	0.189	0.194	0.201	0.204	0.205	0.208	0.212	0.215	0.094
Number of Repeat FITs (total)	0.136	0.138	0.141	0.146	0.148	0.150	0.157	0.159	0.161	0.163	0.168	0.169	0.075
Number of Watch and Wait (total) (total)	0.259	0.259	0.259	0.284	0.284	0.284	0.302	0.302	0.313	0.313	0.331	0.331	0.150
Number of COLs (total)	0.554	0.552	0.549	0.522	0.520	0.518	0.495	0.493	0.481	0.480	0.459	0.457	0.709
Reduction in number of referrals (total - 2WW+18WW)	21.9%	22.2%	22.5%	26.4%	26.7%	27.0%	30.3%	30.5%	32.2%	32.4%	35.4%	35.6%	-
Reduction in number of referrals (2WW only)	36.9%	37.9%	39.0%	43.4%	44.4%	45.5%	50.2%	50.8%	52.9%	53.6%	57.6%	58.2%	-
Increase in number of referrals (18WW only)	87.1%	91.9%	97.0%	97.2%	102.0%	107.1%	114.5%	117.5%	118.8%	121.8%	126.0%	129.0%	-
Increase in number of repeat FITs	81.7%	84.7%	87.9%	94.3%	97.3%	100.5%	109.7%	111.6%	115.0%	116.9%	124.1%	126.0%	-
Increase in number of watch and waits	72.6%	72.6%	72.6%	89.5%	89.5%	89.5%	101.7%	101.7%	108.8%	108.8%	120.9%	120.9%	-
Reduction in number of COLs	21.9%	22.2%	22.5%	26.3%	26.6%	26.9%	30.2%	30.4%	32.1%	32.3%	35.3%	35.5%	-
Mean time to diagnosis - CRC	1.934	1.966	2.031	2.015	2.046	2.108	2.213	2.278	2.308	2.371	2.568	2.629	1.303
Mean time to diagnosis - AAs	3.892	3.956	4.039	4.309	4.371	4.450	4.878	4.931	5.199	5.251	5.730	5.780	1.956
Mean time to diagnosis - IBD	2.405	2.458	2.535	2.660	2.712	2.786	3.046	3.100	3.241	3.294	3.678	3.730	1.684

[□]*Southwest quadrant ICER;* [§] DG30 & NG12

AA, advanced adenomas; COL, colonoscopy; CRC, colorectal cancer; FIT, quantitative faecal immunochemical test; IBD, inflammatory bowel disease; NMB, net monetary benefit; t, threshold

Table 25: Tabulated results for HM-JACKarc using one threshold (Table 94 of the EAG report)

	Int 1: FIT 1 threshold (µg/g)									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.167	14.167	14.167	14.167	14.166	14.166	14.165	14.165	14.164	14.168
QALYs	10.894	10.894	10.894	10.893	10.893	10.893	10.892	10.892	10.892	10.895
Costs (£)	3144	3095	3080	3059	3039	3004	2969	2964	2949	3247
ICER (pairwise, vs Intervention 3)□ (£)	102,358	126,577	129,749	131,152	128,737	119,638	101,797	98,121	88,441	-
NMB λ=20,000 (vs Int 3) (£)	83	128	141	159	176	202	223	225	230	-
NMB λ=30,000 (vs Int 3) (£)	74	117	129	146	161	183	197	198	198	-
Number of 2WW referrals (total)	0.451	0.387	0.368	0.342	0.317	0.276	0.237	0.230	0.215	0.681
Number of 18WW referrals (total)	0.162	0.180	0.186	0.194	0.201	0.213	0.225	0.226	0.231	0.094
Number of Repeat FITs (total)	0.129	0.144	0.149	0.155	0.161	0.170	0.180	0.181	0.185	0.075
Number of Watch and Wait (total) (total)	0.258	0.289	0.297	0.310	0.321	0.340	0.359	0.362	0.370	0.150
Number of COLs (total)	0.560	0.519	0.507	0.490	0.474	0.448	0.422	0.418	0.408	0.709
Reduction in number of referrals (total - 2WW + 18WW)	21.0%	26.8%	28.5%	30.9%	33.2%	36.9%	40.5%	41.1%	42.5%	2.6%
Reduction in number of referrals (2WW only)	33.8%	43.3%	46.0%	49.8%	53.5%	59.4%	65.3%	66.2%	68.5%	0.0%
Increase in number of referrals (18WW only)□□	72.4%	92.5%	98.3%	106.5%	114.4%	127.1%	139.6%	141.6%	146.5%	-
Reduction in number of COLs	20.9%	26.8%	28.4%	30.8%	33.1%	36.8%	40.4%	41.0%	42.4%	-
Mean time to diagnosis - CRC	1.973	2.077	2.123	2.208	2.320	2.611	3.167	3.302	3.720	1.303
Mean time to diagnosis - AAs	3.752	4.215	4.395	4.692	5.047	5.635	6.394	6.540	6.914	1.956
Mean time to diagnosis - IBD	2.318	2.599	2.707	2.889	3.106	3.590	4.291	4.432	4.818	1.684

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 26: Tabulated results for HM-JACKarc using two thresholds (Table 95 of the EAG report)

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.167	14.167	14.167	14.167	14.167	14.167	14.166	14.166	14.166	14.166	14.166	14.166	14.168
QALYs	10.894	10.894	10.894	10.894	10.894	10.894	10.893	10.893	10.893	10.893	10.893	10.893	10.895
Costs (£)	3140	3138	3137	3092	3091	3089	3055	3054	3035	3034	3003	3002	3247
ICER (pairwise, vs Intervention 3) □ (£)	99,786	96,918	91,787	124,149	121,155	115,725	122,828	118,504	122,473	118,660	116,651	113,788	-
NMB λ=20,000 (vs Int 3) (£)	86	86	86	130	130	130	161	160	177	177	202	202	-
NMB λ=30,000 (vs Int 3) (£)	75	75	74	117	117	117	145	144	160	159	181	181	-
Number of 2WW referrals (total)	0.427	0.420	0.413	0.374	0.367	0.360	0.323	0.320	0.303	0.299	0.269	0.266	0.681
Number of 18WW referrals (total)	0.177	0.182	0.187	0.189	0.193	0.198	0.206	0.209	0.210	0.213	0.217	0.220	0.094
Number of Repeat FITs (total)	0.137	0.139	0.142	0.148	0.151	0.153	0.161	0.162	0.165	0.167	0.173	0.174	0.075
Number of Watch and Wait (total) (total)	0.258	0.258	0.258	0.289	0.289	0.289	0.310	0.310	0.321	0.321	0.340	0.340	0.150
Number of COLs (total)	0.553	0.551	0.549	0.515	0.513	0.511	0.485	0.483	0.470	0.469	0.446	0.445	0.709
Reduction in number of referrals (total - 2WW + 18WW)	22.0%	22.3%	22.6%	27.4%	27.7%	28.0%	31.7%	31.9%	33.8%	34.0%	37.2%	37.3%	-
Reduction in number of referrals (2WW only)	37.3%	38.3%	39.4%	45.1%	46.1%	47.1%	52.5%	53.1%	55.6%	56.1%	60.5%	61.0%	-
Increase in number of referrals (18WW only)	89.2%	94.3%	99.3%	101.3%	106.3%	111.4%	119.8%	122.5%	124.5%	127.2%	132.1%	134.9%	-
Increase in number of repeat FITs	82.9%	86.0%	89.2%	98.0%	101.2%	104.3%	114.8%	116.5%	120.7%	122.4%	130.2%	132.0%	-
Increase in number of watch and waits	72.4%	72.4%	72.4%	92.5%	92.5%	92.5%	106.5%	106.5%	114.4%	114.4%	127.1%	127.1%	-
Reduction in number of COLs	21.9%	22.2%	22.5%	27.3%	27.6%	27.9%	31.6%	31.8%	33.7%	33.9%	37.1%	37.3%	-
Mean time to diagnosis - CRC	2.012	2.044	2.104	2.102	2.133	2.191	2.305	2.361	2.404	2.458	2.664	2.717	1.303
Mean time to diagnosis - AAs	3.891	3.955	4.038	4.300	4.361	4.440	4.861	4.914	5.179	5.230	5.707	5.757	1.956
Mean time to diagnosis - IBD	2.404	2.457	2.534	2.652	2.703	2.776	3.031	3.084	3.223	3.276	3.658	3.709	1.684

□*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 27: Tabulated results for IDK Haemoglobin using one threshold (Table 96 of the EAG report)

	Int 1: FIT 1threshold	**Int 3: DG30&NG12**
t (µg/g)	**2**	**10**
LYs	14.166	14.168
QALYs	10.893	10.895
Costs (£)	3012	3247
ICER (pairwise, vs Intervention 3)□ (£)	167,879	-
NMB λ=20,000 (vs Int 3) (£)	206	-
NMB λ=30,000 (vs Int 3) (£)	192	-
Number of 2WW referrals (total)	0.285	0.681
Number of 18WW referrals (total)	0.210	0.094
Number of Repeat FITs (total)	0.168	0.075
Number of Watch and Wait (total) (total)	0.336	0.150
Number of COLs (total)	0.454	0.709
Reduction in number of referrals (total - 2WW + 18WW)	36.1%	-
Reduction in number of referrals (2WW only)	58.1%	-
Increase in number of referrals (18WW only) □□	124.4%	-
Reduction in number of COLs	35.9%	-
Mean time to diagnosis - CRC	2.466	1.303
Mean time to diagnosis - AAs	3.631	1.956
Mean time to diagnosis - IBD	2.199	1.684

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 28: Tabulated results for IDK Haemoglobin/Hapto using one threshold (Table 97 of the EAG report)

	Int 1: FIT 1threshold	**Int 3: DG30&NG12**
t (µg/g)	**2**	**10**
LYs	14.165	14.168
QALYs	10.893	10.895
Costs (£)	3054	3247
ICER (pairwise, vs Intervention 3)□ (£)	112,151	-
NMB λ=20,000 (vs Int 3) (£)	158	-
NMB λ=30,000 (vs Int 3) (£)	141	-
Number of 2WW referrals (total)	0.336	0.681
Number of 18WW referrals (total)	0.195	0.094
Number of Repeat FITs (total)	0.156	0.075
Number of Watch and Wait (total) (total)	0.312	0.150
Number of COLs (total)	0.487	0.709
Reduction in number of referrals (total - 2WW + 18WW)	31.4%	2.6%
Reduction in number of referrals (2WW only)	50.6%	0.0%
Increase in number of referrals (18WW only) □□	108.3%	-
Reduction in number of COLs	31.3%	-
Mean time to diagnosis - CRC	2.791	1.303
Mean time to diagnosis - AAs	3.668	1.956
Mean time to diagnosis - IBD	2.236	1.684

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*
AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 29: Tabulated results for NS Prime (Table 98 of the EAG report)

	Int 1: FIT 1 threshold			Int 2: FIT 2 thresholds			Int 3: DG30&NG12
Threshold - t or t_{low}/t_{high} (µg/g)	**3**	**10**	**100**	**3/10**	**3/100**	**10/100**	**10**
LYs	14.165	14.164	14.163	14.165	14.165	14.164	14.168
QALYs	10.893	10.892	10.891	10.893	10.893	10.892	10.895
Costs (£)	3,303	3,029	2,933	3,297	3,296	3,025	3247
ICER (pairwise, vs Intervention 3)[□] (£)	Dominated	80,090	79,835	Dominated	Dominated	78,286	-
NMB λ=20,000 (vs Int 3) (£)	-96	163	235	-92	-93	165	-
NMB λ=30,000 (vs Int 3) (£)	-115	136	196	-113	-116	137	-
Number of 2WW referrals (total)	0.683	0.307	0.196	0.616	0.597	0.288	0.681
Number of 18WW referrals (total)	0.093	0.204	0.237	0.138	0.151	0.217	0.094
Number of Repeat FITs (total)	0.075	0.163	0.189	0.097	0.103	0.170	0.075
Number of Watch and Wait (total) (total)	0.149	0.326	0.379	0.149	0.149	0.326	0.150
Number of COLs (total)	0.709	0.468	0.396	0.689	0.683	0.462	0.709
Reduction in number of referrals (total - 2WW + 18WW)	-0.1%	34.1%	44.2%	2.7%	3.6%	34.9%	-
Reduction in number of referrals (2WW)	-0.2%	54.9%	71.3%	9.5%	12.4%	57.8%	-
Increase in number of referrals (18WW)	-0.4%	117.4%	152.5%	46.7%	60.8%	131.4%	-
Increase in number of repeat FITs	-0.4%	117.4%	152.5%	29.1%	37.8%	126.2%	-
Increase in number of watch and waits	-0.4%	117.4%	152.5%	-0.4%	-0.4%	117.4%	-
Reduction in number of COLs	-0.0%	34.0%	44.2%	2.8%	3.6%	34.8%	-
Mean time to diagnosis - CRC	2.828	3.515	4.348	2.922	3.031	3.598	1.303
Mean time to diagnosis - AAs	4.220	5.038	6.888	4.332	4.568	5.218	1.956
Mean time to diagnosis - IBD	2.676	3.097	4.795	2.738	2.954	3.265	1.684

[□]*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 30: Tabulated results for OC Sensor using one threshold (Table 99 of the EAG report)

	Int 1: FIT 1 threshold (µg/g)									Int 3 : DG30& NG12
t (µg/g)	**2**	**4**	**5**	**7**	**10**	**20**	**50**	**60**	**100**	**10**
LYs	14.167	14.167	14.167	14.167	14.166	14.166	14.165	14.164	14.164	14.168
QALYs	10.894	10.894	10.894	10.893	10.893	10.893	10.892	10.892	10.891	10.895
Costs (£)	3224	3155	3133	3101	3069	3018	2970	2963	2946	3247
ICER (pairwise, vs Intervention 3)□ (£)	20,918	73,459	85,983	99,929	108,258	111,840	99,748	96,391	87,017	-
NMB λ=20,000 (vs Int 3) (£)	1	67	87	117	145	188	221	225	232	-
NMB λ=30,000 (vs Int 3) (£)	-10	54	74	102	128	167	193	195	197	-
Number of 2WW referrals (total)	0.561	0.465	0.435	0.394	0.354	0.292	0.237	0.229	0.210	0.681
Number of 18WW referrals (total)	0.129	0.157	0.166	0.178	0.190	0.208	0.224	0.227	0.232	0.094
Number of Repeat FITs (total)	0.103	0.126	0.133	0.143	0.152	0.167	0.180	0.181	0.186	0.075
Number of Watch and Wait (total) (total)	0.206	0.252	0.266	0.285	0.304	0.333	0.359	0.363	0.372	0.150
Number of COLs (total)	0.631	0.569	0.550	0.523	0.498	0.458	0.422	0.417	0.405	0.709
Reduction in number of referrals (total - 2WW + 18WW)	10.9%	19.7%	22.4%	26.2%	29.8%	35.4%	40.5%	41.2%	42.9%	-
Reduction in number of referrals (2WW only)	17.6%	31.8%	36.1%	42.2%	48.0%	57.1%	65.2%	66.4%	69.1%	-
Increase in number of referrals (18WW only)□□	37.7%	68.1%	77.3%	90.3%	102.7%	122.2%	139.5%	142.0%	147.9%	-
Reduction in number of COLs	10.9%	19.7%	22.4%	26.1%	29.7%	35.4%	40.4%	41.1%	42.8%	-
Mean time to diagnosis - CRC	2.008	2.094	2.136	2.215	2.324	2.621	3.219	3.368	3.827	1.303
Mean time to diagnosis - AAs	3.833	4.279	4.453	4.739	5.084	5.653	6.394	6.538	6.908	1.956
Mean time to diagnosis - IBD	2.397	2.659	2.761	2.932	3.138	3.606	4.291	4.431	4.813	1.684

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 31: Tabulated results for OCSensor using two thresholds (Table 100 of the EAG report)

	Int 2: FIT 2 thresholds												Int 3: DG30& NG12
t_{low}/t_{high} (µg/g)	**2/10**	**2/20**	**2/50**	**4/10**	**4/20**	**4/50**	**7/50**	**7/100**	**10/50**	**10/100**	**20/50**	**20/100**	**10**
LYs	14.167	14.167	14.167	14.167	14.167	14.167	14.166	14.166	14.166	14.166	14.166	14.166	14.168
QALYs	10.894	10.894	10.894	10.894	10.894	10.893	10.893	10.893	10.893	10.893	10.893	10.893	10.895
Costs (£)	3219	3218	3216	3152	3150	3148	3095	3094	3065	3064	3016	3015	3247
ICER (pairwise, vs Intervention 3)□ (£)	23,951	24,217	23,657	73,581	72,507	69,597	94,780	91,290	103,677	100,282	109,219	106,415	-
NMB λ=20,000 (vs Int 3) (£)	5	5	5	69	70	70	120	119	147	146	189	189	-
NMB λ=30,000 (vs Int 3) (£)	-7	-7	-8	56	57	56	104	102	129	128	167	167	-
Number of 2WW referrals (total)	0.525	0.514	0.504	0.445	0.434	0.424	0.366	0.361	0.333	0.329	0.282	0.278	0.681
Number of 18WW referrals (total)	0.153	0.161	0.167	0.170	0.178	0.184	0.197	0.200	0.204	0.207	0.215	0.218	0.094
Number of Repeat FITs (total)	0.115	0.119	0.122	0.132	0.136	0.139	0.152	0.153	0.159	0.160	0.170	0.171	0.075
Number of Watch and Wait (total) (total)	0.206	0.206	0.206	0.252	0.252	0.252	0.285	0.285	0.304	0.304	0.333	0.333	0.150
Number of COLs (total)	0.620	0.617	0.614	0.563	0.560	0.557	0.515	0.514	0.492	0.490	0.455	0.454	0.709
Reduction in number of referrals (total - 2WW + 18WW)	12.5%	13.0%	13.4%	20.6%	21.1%	21.5%	27.4%	27.6%	30.7%	30.9%	35.9%	36.1%	-
Reduction in number of referrals (2WW only)	23.0%	24.6%	26.0%	34.7%	36.3%	37.7%	46.3%	47.0%	51.1%	51.8%	58.5%	59.2%	-
Increase in number of referrals (18WW only)	63.7%	71.5%	78.4%	81.9%	89.7%	96.6%	110.0%	113.3%	117.4%	120.8%	129.1%	132.4%	-
Increase in number of repeat FITs	53.9%	58.8%	63.1%	76.7%	81.6%	85.9%	102.6%	104.7%	111.9%	114.0%	126.5%	128.6%	-
Increase in number of watch and waits	37.7%	37.7%	37.7%	68.1%	68.1%	68.1%	90.3%	90.3%	102.7%	102.7%	122.2%	122.2%	-
Reduction in number of COLs	12.5%	13.0%	13.4%	20.5%	21.0%	21.4%	27.3%	27.5%	30.6%	30.8%	35.8%	36.0%	-
Mean time to diagnosis - CRC	2.048	2.083	2.154	2.120	2.153	2.219	2.321	2.385	2.415	2.477	2.679	2.739	1.303
Mean time to diagnosis - AAs	3.979	4.046	4.135	4.367	4.430	4.512	4.912	4.966	5.217	5.269	5.724	5.774	1.956
Mean time to diagnosis - IBD	2.486	2.542	2.623	2.713	2.765	2.841	3.075	3.131	3.256	3.310	3.673	3.724	1.684

□*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold;

Table 32: Tabulated results for QuikRead go (Table 101 of the EAG report)

	Int 1: FIT 1 threshold			Int 2: FIT 2 thresholds			Int 3: DG30& NG12
t (µg/g)	**10**	**100**	**150**	**10/100**	**10/150**	**100/150**	**10**
LYs	14.167	14.164	14.163	14.166	14.166	14.164	14.168
QALYs	3112	2953	2939	3106	3105	2953	3247
Costs (£)	89,603	94,538	76,730	84,390	80,194	92,330	-
ICER (pairwise, vs Intervention 3)□ (£)	104	231	228	108	106	230	-
NMB λ=20,000 (vs Int 3) (£)	103	230	227	107	105	229	-
NMB λ=30,000 (vs Int 3) (£)	88	199	187	90	88	198	-
Number of 2WW referrals (total)	0.408	0.218	0.204	0.374	0.372	0.215	0.681
Number of 18WW referrals (total)	0.174	0.230	0.234	0.197	0.198	0.232	0.094
Number of Repeat FITs (total)	0.139	0.184	0.187	0.151	0.151	0.185	0.075
Number of Watch and Wait (total)	0.279	0.368	0.375	0.279	0.279	0.368	0.150
Number of COLs (total)	0.532	0.410	0.401	0.522	0.521	0.409	0.709
Reduction in number of referrals (total - 2WW + 18WW)	24.9%	42.2%	43.5%	26.4%	26.5%	42.3%	-
Reduction in number of referrals (2WW only)	40.2%	68.0%	70.1%	45.1%	45.4%	68.4%	-
Increase in number of referrals (18WW only)	85.9%	145.5%	149.9%	109.7%	111.5%	147.3%	-
Increase in number of repeat FITs	85.9%	145.5%	149.9%	100.8%	101.9%	146.6%	-
Increase in number of watch and waits	85.9%	145.5%	149.9%	85.9%	85.9%	145.5%	-
Reduction in number of COLs	24.9%	42.1%	43.4%	26.3%	26.4%	42.3%	-
Mean time to diagnosis - CRC	2.148	3.436	4.356	2.285	2.382	3.521	1.303
Mean time to diagnosis - AAs	5.137	6.919	7.172	5.327	5.354	6.942	1.956
Mean time to diagnosis - IBD	3.185	4.822	5.116	3.360	3.392	4.849	1.684

□*Southwest quadrant ICER*

AA – advanced adenomas; COL - colonoscopy; CRC – colorectal cancer; FIT – quantitative faecal immunochemical test; IBD - inflammatory bowel disease; NMB – net monetary benefit; t – threshold

5. Additional scenarios ran by the EAG

Additional scenario 1: Reduction in prevalence for CRC, AAs and IBD by 50%

Table 33: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.244	14.244	14.244	14.244	14.244	14.243	14.243	14.242	14.242	14.245
QALYs	11.030	11.030	11.030	11.030	11.030	11.029	11.029	11.029	11.028	11.031
Costs (£)	1773	1703	1683	1654	1627	1582	1539	1533	1516	1981
ICER (pairwise, vs Intervention 3)□ (£)	321,847	356,839	356,459	348,151	331,020	292,842	237,002	226,528	199,903	-
iNMB λ=20,000 (vs Int 3) (£)	196	263	282	309	333	372	405	409	419	-
iNMB λ=30,000 (vs Int 3) (£)	189	255	274	299	322	358	386	389	395	-
Number of 2WW referrals (total)	0.368	0.295	0.274	0.245	0.217	0.173	0.130	0.124	0.108	0.640
Number of 18WW referrals (total)	0.066	0.074	0.076	0.079	0.082	0.087	0.092	0.092	0.094	0.038
Number of Repeat FITs (total)	0.066	0.074	0.076	0.079	0.082	0.087	0.092	0.092	0.094	0.038
Number of Watch and Wait (total) (total)	0.499	0.556	0.573	0.596	0.618	0.653	0.687	0.692	0.704	0.284
Number of COLs (total)	0.397	0.338	0.321	0.297	0.274	0.238	0.203	0.198	0.185	0.618
Reduction in number of referrals (total - 2WW + 18WW)	35.9%	45.5%	48.3%	52.1%	55.8%	61.7%	67.3%	68.1%	70.2%	-
Reduction in number of referrals (2WW only)	42.4%	53.9%	57.1%	61.7%	66.0%	73.0%	79.6%	80.7%	83.2%	-
Increase in number of referrals (18WW only)□□	75.5%	95.8%	101.6%	109.8%	117.5%	129.9%	141.7%	143.5%	148.0%	-
Reduction in number of COLs	35.8%	45.4%	48.2%	52.0%	55.7%	61.6%	67.2%	68.1%	70.2%	-
Mean time to diagnosis - CRC	2.298	2.448	2.516	2.639	2.802	3.224	4.033	4.230	4.838	1.384
Mean time to diagnosis - AAs	4.443	5.114	5.376	5.807	6.326	7.192	8.315	8.532	9.089	1.956
Mean time to diagnosis - IBD	2.933	3.341	3.498	3.762	4.077	4.783	5.807	6.014	6.579	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Additional scenario 2: Increase in prevalence for CRC, AAs and IBD by 50%

Table 34: Tabulated results for HM JACKarc using one threshold

	Int 1: FIT 1 threshold									Int 3 : DG30& NG12
t (µg/g)	2	4	5	7	10	20	50	60	100	10
LYs	14.089	14.088	14.088	14.088	14.087	14.086	14.084	14.083	14.082	14.091
QALYs	10.757	10.756	10.756	10.756	10.756	10.755	10.753	10.753	10.752	10.759
Costs (£)	4136	4075	4057	4031	4006	3964	3920	3912	3893	4304
ICER (pairwise, vs Intervention 3)□ (£)	78,664	89,032	89,604	88,583	85,416	77,549	64,973	62,524	56,193	-
iNMB λ=20,000 (vs Int 3) (£)	125	178	192	211	228	252	266	266	264	-
iNMB λ=30,000 (vs Int 3) (£)	104	152	164	181	193	209	207	204	191	-
Number of 2WW referrals (total)	0.404	0.334	0.313	0.284	0.256	0.210	0.163	0.155	0.136	0.648
Number of 18WW referrals (total)	0.063	0.070	0.072	0.075	0.078	0.083	0.088	0.089	0.091	0.037
Number of Repeat FITs (total)	0.063	0.070	0.072	0.075	0.078	0.083	0.088	0.089	0.091	0.037
Number of Watch and Wait (total) (total)	0.470	0.526	0.542	0.565	0.587	0.624	0.661	0.667	0.682	0.278
Number of COLs (total)	0.430	0.372	0.355	0.332	0.309	0.271	0.232	0.226	0.210	0.629
Reduction in number of referrals (total - 2WW + 18WW)	31.8%	41.0%	43.7%	47.5%	51.2%	57.2%	63.3%	64.3%	66.8%	-
Reduction in number of referrals (2WW only)	37.6%	48.5%	51.7%	56.1%	60.5%	67.6%	74.8%	76.0%	78.9%	-
Increase in number of referrals (18WW only)□□	-69.1%	-89.1%	-94.9%	-103.2%	-111.2%	-124.3%	-137.5%	-139.7%	-145.1%	-
Reduction in number of COLs	31.6%	40.8%	43.5%	47.3%	50.9%	56.9%	63.1%	64.1%	66.6%	-
Mean time to diagnosis - CRC	2.317	2.471	2.539	2.663	2.825	3.247	4.055	4.251	4.858	1.384
Mean time to diagnosis - AAs	4.464	5.138	5.401	5.833	6.353	7.220	8.341	8.558	9.112	1.956
Mean time to diagnosis - IBD	2.955	3.365	3.523	3.788	4.104	4.810	5.832	6.038	6.602	2.044

□*Southwest quadrant ICER;* □□ *Also the value for increased repeat FITs and increased number of watch and waits*

AAs – Advanced adenomas; FIT - faecal immunochemical test; IBD – inflammatory bowel disease; iNMB – incremental net monetary benefit; LY – life years; QALY - quality-adjusted life year

Addendum 2 to the Technology Assessment Report commissioned by the NIHR Evidence Synthesis Programme on behalf of the National Institute for Health and Care Excellence – Diagnostics Assessment Report Guide

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

Produced by	School of Health and Related Research (ScHARR), The University of Sheffield
Authors	Sue Harnan, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Kate (Shijie) Ren, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Jean Hamilton, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Emma Simpson, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Aline Navega Biz, Research Fellow, ScHARR, University of Sheffield, Sheffield, UK Matt Stevenson, Professor of Health Technology Assessment, ScHARR, University of Sheffield, Sheffield, UK
Correspondence Author	Sue Harnan, Senior Research Fellow, ScHARR, University of Sheffield, Sheffield, UK
Date completed	21/06/2023

Source of funding: This report was commissioned by the NIHR Evidence Synthesis Programme as project number NIHR135637.

1. Introduction

This addendum contains all data used in the statistical syntheses of diagnostic test accuracy.

Please note that there may be some differences in the sensitivity and specificity reported here and in the report. This is because the data in the report is as extracted from the source papers, whereas this data has been calculated from the TP, TN, FP, FN that was used in the synthesis. In order to include studies that did not report TP, TN etc in the synthesis, these values were sometimes back-calculated from other available data, e.g. sensitivity, specificity, total patients in the analysis, number of CRC events in the analysis and where numbers were large, there were multiple plausible values for TP, TN etc. (see Appendix 2).

Please also note that the EAG carefully selected studies to enter each analysis and sensitivity analysis to avoid double counting of patients who may have been included in multiple analyses in the evidence base (see Section 4.1.7 of the main EAG report).

2. Data presented

Data relating to the diagnostic test accuracy for the detection of CRC by all tests are provided in Table 1. Data relating to the diagnostic test accuracy for the detection of advanced adenomas and inflammatory bowel disease by all tests are provided in **Table *2***.

Table 1 Data entering the statistical syntheses of the diagnostic test accuracy of FITs for detection of CRC

Author, year	Test	Out-come	Pop. type	Sub-group	Ref Stand	No. Pts	CRC cases	Prev CRC	Thres-hold (µg/g)	TP	TN	FN	FP	Sensitivity	Specificity
FOB gold															
Benton 2022[1]	FOB Gold Wide - SENTiFIT 270	CRC	2	0	1	233	7	0.030043	2	4	186	3	40	57.1 (50.7,63.5)	82.3 (77.4,87.2)
Benton 2022[1]	FOB Gold Wide - SENTiFIT 270	CRC	2	0	1	233	7	0.030043	10	4	211	3	15	57.1 (50.7,63.5)	93.4 (90.2,96.6)
Benton 2022[1]	FOB Gold Wide - SENTiFIT 270	CRC	2	0	1	233	7	0.030043	100	4	219	3	7	57.1 (50.7,63.5)	96.9 (94.7,99.1)
MacLean 2022a[2]	FOB Gold Wide TÂ® with SENTiFITÂ® 270 analyser	CRC	2	0	1	553	14	0.025316	3	14	415	0	124	100 (NE,NE)	77 (73.5,80.5)
MacLean 2022a[2]	FOB Gold Wide TÂ® with SENTiFITÂ® 270 analyser	CRC	2	0	1	553	14	0.025316	10	14	457	0	82	100 (NE,NE)	84.8 (81.8,87.8)
MacLean 2022a[2]	FOB Gold Wide TÂ® with SENTiFITÂ® 270 analyser	CRC	2	0	1	553	14	0.025316	100	13	506	1	33	92.9 (90.8,95)	93.9 (91.9,95.9)
MacLean 2022a[2]	FOB Gold Wide TÂ® with SENTiFITÂ® 270 analyser	CRC	2	0	1	553	14	0.025316	150	11	511	3	28	78.6 (75.2,82)	94.8 (92.9,96.7)
Schwettmann 2022[3]	FOB Gold + Roche Cobas 8000 c702 analyser	CRC	4	0	1	163	26	0.159509	10	25	71	1	66	96.2 (93.3,99.1)	51.8 (44.1,59.5)
Schwettmann 2022[3]	FOB Gold + Roche Cobas 8000 c702 analyser	CRC	4	0	1	163	26	0.159509	15	25	79	1	58	96.2 (93.3,99.1)	57.7 (50.1,65.3)
Schwettmann 2022[3]	FOB Gold + Roche Cobas 8000 c702 analyser	CRC	4	0	1	163	26	0.159509	20	25	83	1	54	96.2 (93.3,99.1)	60.6 (53.1,68.1)
Schwettmann 2022[3]	FOB Gold + Roche Cobas 8000 c702 analyser	CRC	4	0	1	163	26	0.159509	30	24	92	2	45	92.3 (88.2,96.4)	67.2 (60,74.4)
Schwettmann 2022[3]	FOB Gold + Roche Cobas 8000 c702 analyser	CRC	4	0	1	163	26	0.159509	40	23	95	3	42	88.5 (83.6,93.4)	69.3 (62.2,76.4)

Schwettmann 2022[3]	FOB Gold + Roche Cobas 8000 c702 analyser	CRC	4	0	1	163	26	0.159509	50	23	99	3	38	88.5 (83.6,93.4)	72.3 (65.4,79.2)
Schwettmann 2022[3]	FOB Gold + Roche Cobas 8000 c702 analyser	CRC	4	0	1	163	26	0.159509	100	21	104	5	33	80.8 (74.8,86.8)	75.9 (69.3,82.5)
Schwettmann 2022[3]	FOB Gold + Roche Cobas 8000 c702 analyser	CRC	4	0	1	163	26	0.159509	150	15	114	11	23	57.7 (50.1,65.3)	83.2 (77.5,88.9)
HM-JACKarc															
Benton 2022[1]	HM-JACKarc	CRC	2	0	1	233	7	0.030043	2	4	154	3	72	57.1 (50.7,63.5)	68.1 (62.1,74.1)
Benton 2022[1]	HM-JACKarc	CRC	2	0	1	233	7	0.030043	10	4	191	3	35	57.1 (50.7,63.5)	84.5 (79.9,89.1)
Benton 2022[1]	HM-JACKarc	CRC	2	0	1	233	7	0.030043	100	4	220	3	6	57.1 (50.7,63.5)	97.3 (95.2,99.4)
Chapman 2021[4]	HM JACKarc + HM JACKarc analyser	CRC	4	0	1	732	38	0.051913	4	35	486	3	208	92 (90,94)	70 (66.7,73.3)
Chapman 2021[4]	HM JACKarc + HM JACKarc analyser	CRC	4	0	1	732	38	0.051913	10	32	541	6	153	84 (81.3,86.7)	78 (75,81)
Chapman 2021[4]	HM JACKarc + HM JACKarc analyser	CRC	4	0	1	732	38	0.051913	22.6	31	562	7	132	82 (79.2,84.8)	81 (78.2,83.8)
Chapman 2021[4]	HM JACKarc + HM JACKarc analyser	CRC	4	0	1	732	38	0.051913	150	22	659	16	35	58 (54.4,61.6)	95 (93.4,96.6)
Cunin 2020[5]	HM JACKarc (analyser NR)	CRC	4	1 - IDA	2	189	20	0.10582	10	16	138	4	31	80 (74.3,85.7)	81.6 (76.1,87.1)
Cunin 2020[5]	HM JACKarc (analyser NR)	CRC	4	1 - no IDA	2	739	28	0.037889	10	25	597	3	114	89 (86.7,91.3)	84 (81.4,86.6)
D'Souza 2020a[6]	HM JACKarc analytical system	CRC	1	0	1	298	12	0.040268	2	12	218	0	68	100 (NE,NE)	76.2 (71.4,81)
D'Souza 2020a[6]	HM JACKarc analytical system	CRC	1	0	1	298	12	0.040268	10	11	253	1	33	92 (88.9,95.1)	88.5 (84.9,92.1)
D'Souza 2020a[6]	HM JACKarc analytical system	CRC	2	0	1	160	8	0.05	2	8	108	0	44	100 (NE,NE)	71 (64,78)

D'Souza 2020a[6]	HM JACKarc analytical system	CRC	2	0	1	160	8	0.05	10	7	128	1	24	87.5 (82.4,92.6)	84.2 (78.5,89.9)
D'Souza 2020a[6]	HM JACKarc analytical system	CRC	3	0	1	138	4	0.028986	2	4	110	0	24	100 (NE,NE)	82.1 (75.7,88.5)
D'Souza 2020a[6]	HM JACKarc analytical system	CRC	3	0	1	138	4	0.028986	10	4	125	0	9	100 (NE,NE)	93.3 (89.1,97.5)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	2	0	1	7194	257	0.035724	2	251	4368	6	2569	97.7 (97.4,98)	63 (61.9,64.1)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	4	0	1	1994	53	0.02658	2	50	1346	3	545	94.3 (93.3,95.3)	71.8 (69.8,73.8)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	4	0	1	634	19	0.029968	2	18	393	1	222	94.7 (93,96.4)	63.9 (60.2,67.6)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	2	0	1	7194	257	0.035724	10	237	5711	20	1226	92.2 (91.6,92.8)	82.3 (81.4,83.2)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	4	0	1	1994	53	0.02658	10	46	1665	7	226	86.8 (85.3,88.3)	88.4 (87,89.8)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	4	0	1	634	19	0.029968	10	16	508	3	107	84.2 (81.4,87)	82.6 (79.6,85.6)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	4	1	1	479	15.807	0.033	10	16	378	0	85	100 (NE,NE)	81.6 (78.1,85.1)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	2	0	1	7194	257	0.035724	150	185	6514	72	423	72 (71,73)	93.9 (93.3,94.5)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	4	0	1	1994	53	0.02658	150	33	1881	20	60	62.3 (60.2,64.4)	96.9 (96.1,97.7)
D'Souza 2021a[7]	HM JACKarc analytical system	CRC	4	0	1	634	19	0.029968	150	15	580	4	35	78.9 (75.7,82.1)	94.3 (92.5,96.1)
D'Souza 2021c[8]	HM JACKarc analytical system	CRC	4	0	1	9822	329	0.033496	2	319	6157	10	3336	97 (96.7,97.3)	64.9 (64,65.8)
D'Souza 2021c[8]	HM JACKarc analytical system	CRC	4	0	1	9822	329	0.033496	10	299	7930	30	1563	90.9 (90.3,91.5)	83.5 (82.8,84.2)
D'Souza 2021c[8]	HM JACKarc analytical system	CRC	4	0	1	9822	329	0.033496	150	233	8977	96	516	70.8 (69.9,71.7)	94.6 (94.2,95)

D'Souza 2021b[9]	HM-JACKarc analyser	CRC	4	7 age <50	1	1103	16	0.014506	2	14	765	2	322	87.5 (85.5,89.5)	70.4 (67.7,73.1)
D'Souza 2021b[9]	HM-JACKarc analyser	CRC	4	7 age 50+	1	8719	313	0.035899	2	305	5388	8	3018	97.4 (97.1,97.7)	64.1 (63.1,65.1)
D'Souza 2021b[9]	HM-JACKarc analyser	CRC	4	7 age <50	1	1103	16	0.014506	10	13	909	3	178	81.3 (79,83.6)	83.6 (81.4,85.8)
D'Souza 2021b[9]	HM-JACKarc analyser	CRC	4	7 age 50+	1	8719	313	0.035899	10	286	7019	27	1387	91.4 (90.8,92)	83.5 (82.7,84.3)
D'Souza 2021b[9]	HM-JACKarc analyser	CRC	4	7 age <50	1	1103	16	0.014506	50	11	1002	5	85	68.8 (66.1,71.5)	92.2 (90.6,93.8)
D'Souza 2021b[9]	HM-JACKarc analyser	CRC	4	7 age 50+	1	8719	313	0.035899	50	222	7977	91	429	70.9 (69.9,71.9)	94.9 (94.4,95.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	6	52	0	0	940	100 (NE,NE)	0 (NE,NE)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	7.5	49	707	3	233	94.2 (92.7,95.7)	75.2 (72.5,77.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	8.5	48	713	4	227	92.3 (90.6,94)	75.9 (73.2,78.6)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	9.5	48	717	4	223	92.3 (90.6,94)	76.3 (73.7,78.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	10.5	48	725	4	215	92.3 (90.6,94)	77.1 (74.5,79.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	11.5	46	730	6	210	88.5 (86.5,90.5)	77.7 (75.1,80.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	12.5	46	735	6	205	88.5 (86.5,90.5)	78.2 (75.6,80.8)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	13.5	46	746	6	194	88.5 (86.5,90.5)	79.4 (76.9,81.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	14.5	46	747	6	193	88.5 (86.5,90.5)	79.5 (77,82)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	15.5	46	750	6	190	88.5 (86.5,90.5)	79.8 (77.3,82.3)

Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	16.5	46	757	6	183	88.5 (86.5,90.5)	80.5 (78,83)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	17.5	46	760	6	180	88.5 (86.5,90.5)	80.9 (78.5,83.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	18.5	46	762	6	178	88.5 (86.5,90.5)	81.1 (78.7,83.5)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	19.5	46	766	6	174	88.5 (86.5,90.5)	81.5 (79.1,83.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	20.5	45	770	7	170	86.5 (84.4,88.6)	81.9 (79.5,84.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	21.5	45	773	7	167	86.5 (84.4,88.6)	82.2 (79.8,84.6)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	22.5	45	776	7	164	86.5 (84.4,88.6)	82.6 (80.2,85)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	23.5	45	777	7	163	86.5 (84.4,88.6)	82.7 (80.3,85.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	24.5	45	778	7	162	86.5 (84.4,88.6)	82.8 (80.5,85.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	25.5	45	781	7	159	86.5 (84.4,88.6)	83.1 (80.8,85.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	26.5	45	783	7	157	86.5 (84.4,88.6)	83.3 (81,85.6)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	27.5	44	787	8	153	84.6 (82.4,86.8)	83.7 (81.4,86)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	28.5	44	788	8	152	84.6 (82.4,86.8)	83.8 (81.5,86.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	29.5	43	788	9	152	82.7 (80.3,85.1)	83.8 (81.5,86.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	30.5	43	790	9	150	82.7 (80.3,85.1)	84 (81.7,86.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	31.5	43	793	9	147	82.7 (80.3,85.1)	84.4 (82.1,86.7)

Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	32.5	43	795	9	145	82.7 (80.3,85.1)	84.6 (82.4,86.8)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	33.5	43	797	9	143	82.7 (80.3,85.1)	84.8 (82.6,87)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	34.5	42	798	10	142	80.8 (78.3,83.3)	84.9 (82.7,87.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	35.5	42	800	10	140	80.8 (78.3,83.3)	85.1 (82.9,87.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	36.5	42	804	10	136	80.8 (78.3,83.3)	85.5 (83.3,87.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	38.5	42	806	10	134	80.8 (78.3,83.3)	85.7 (83.5,87.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	40.5	41	809	11	131	78.8 (76.3,81.3)	86.1 (83.9,88.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	41.5	41	811	11	129	78.8 (76.3,81.3)	86.3 (84.2,88.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	42.5	40	814	12	126	76.9 (74.3,79.5)	86.6 (84.5,88.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	44	40	815	12	125	76.9 (74.3,79.5)	86.7 (84.6,88.8)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	46	40	816	12	124	76.9 (74.3,79.5)	86.8 (84.7,88.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	47.5	39	818	13	122	75 (72.3,77.7)	87 (84.9,89.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	48.5	39	819	13	121	75 (72.3,77.7)	87.1 (85,89.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	50	39	822	13	118	75 (72.3,77.7)	87.4 (85.3,89.5)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	51.5	39	824	13	116	75 (72.3,77.7)	87.7 (85.7,89.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	52.5	39	829	13	111	75 (72.3,77.7)	88.2 (86.2,90.2)

Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	53.5	39	830	13	110	75 (72.3,77.7)	88.3 (86.3,90.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	54.5	38	833	14	107	73.1 (70.3,75.9)	88.6 (86.6,90.6)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	55.5	38	834	14	106	73.1 (70.3,75.9)	88.7 (86.7,90.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	56.5	37	835	15	105	71.2 (68.4,74)	88.8 (86.8,90.8)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	57.5	36	837	16	103	69.2 (66.3,72.1)	89 (87.1,90.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	58.5	36	839	16	101	69.2 (66.3,72.1)	89.3 (87.4,91.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	59.5	35	840	17	100	67.3 (64.4,70.2)	89.4 (87.5,91.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	60.5	34	841	18	99	65.4 (62.4,68.4)	89.5 (87.6,91.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	61.5	33	841	19	99	63.5 (60.5,66.5)	89.5 (87.6,91.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	62.5	33	842	19	98	63.5 (60.5,66.5)	89.6 (87.7,91.5)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	63.5	33	843	19	97	63.5 (60.5,66.5)	89.7 (87.8,91.6)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	65	33	844	19	96	63.5 (60.5,66.5)	89.8 (87.9,91.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	66.5	32	844	20	96	61.5 (58.5,64.5)	89.8 (87.9,91.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	67.5	32	845	20	95	61.5 (58.5,64.5)	89.9 (88,91.8)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	68.5	32	846	20	94	61.5 (58.5,64.5)	90 (88.1,91.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	69.5	32	847	20	93	61.5 (58.5,64.5)	90.1 (88.2,92)

Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	71	31	848	21	92	59.6 (56.5,62.7)	90.2 (88.3,92.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	74	31	849	21	91	59.6 (56.5,62.7)	90.3 (88.5,92.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	76.5	31	850	21	90	59.6 (56.5,62.7)	90.4 (88.6,92.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	78.5	31	851	21	89	59.6 (56.5,62.7)	90.5 (88.7,92.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	81	31	853	21	87	59.6 (56.5,62.7)	90.7 (88.9,92.5)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	83	30	854	22	86	57.7 (54.6,60.8)	90.9 (89.1,92.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	85	30	855	22	85	57.7 (54.6,60.8)	91 (89.2,92.8)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	87	30	857	22	83	57.7 (54.6,60.8)	91.2 (89.4,93)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	89	30	858	22	82	57.7 (54.6,60.8)	91.3 (89.5,93.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	92.5	29	858	23	82	55.8 (52.7,58.9)	91.3 (89.5,93.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	96.5	29	859	23	81	55.8 (52.7,58.9)	91.4 (89.7,93.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	100	29	860	23	80	55.8 (52.7,58.9)	91.5 (89.8,93.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	106	29	861	23	79	55.8 (52.7,58.9)	91.6 (89.9,93.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	110.5	28	862	24	78	53.8 (50.7,56.9)	91.7 (90,93.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	111.5	28	864	24	76	53.8 (50.7,56.9)	91.9 (90.2,93.6)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	114	28	865	24	75	53.8 (50.7,56.9)	92 (90.3,93.7)

Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	116.5	28	867	24	73	53.8 (50.7,56.9)	92.2 (90.5,93.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	118	28	868	24	72	53.8 (50.7,56.9)	92.3 (90.6,94)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	120.5	28	869	24	71	53.8 (50.7,56.9)	92.4 (90.8,94)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	123.5	28	870	24	70	53.8 (50.7,56.9)	92.6 (91,94.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	126.5	27	870	25	70	51.9 (48.8,55)	92.6 (91,94.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	132.5	26	870	26	70	50 (46.9,53.1)	92.6 (91,94.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	139	25	870	27	70	48.1 (45,51.2)	92.6 (91,94.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	150.5	25	871	27	69	48.1 (45,51.2)	92.7 (91.1,94.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	162	25	872	27	68	48.1 (45,51.2)	92.8 (91.2,94.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	165	24	872	28	68	46.2 (43.1,49.3)	92.8 (91.2,94.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	167.5	24	873	28	67	46.2 (43.1,49.3)	92.9 (91.3,94.5)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	195.5	24	875	28	65	46.2 (43.1,49.3)	93.1 (91.5,94.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	222.5	24	876	28	64	46.2 (43.1,49.3)	93.2 (91.6,94.8)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	227	24	877	28	63	46.2 (43.1,49.3)	93.3 (91.7,94.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	241.5	24	878	28	62	46.2 (43.1,49.3)	93.4 (91.9,94.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	260.5	23	878	29	62	44.2 (41.1,47.3)	93.4 (91.9,94.9)

Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	276	23	879	29	61	44.2 (41.1,47.3)	93.5 (92,95)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	285	23	880	29	60	44.2 (41.1,47.3)	93.6 (92.1,95.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	292	23	881	29	59	44.2 (41.1,47.3)	93.7 (92.2,95.2)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	297.5	23	882	29	58	44.2 (41.1,47.3)	93.8 (92.3,95.3)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	303	23	883	29	57	44.2 (41.1,47.3)	93.9 (92.4,95.4)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	308.5	23	885	29	55	44.2 (41.1,47.3)	94.1 (92.6,95.6)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	322	23	886	29	54	44.2 (41.1,47.3)	94.3 (92.9,95.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	348.5	22	886	30	54	42.3 (39.2,45.4)	94.3 (92.9,95.7)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	373	22	887	30	53	42.3 (39.2,45.4)	94.4 (93,95.8)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	387.5	22	888	30	52	42.3 (39.2,45.4)	94.5 (93.1,95.9)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	395	22	889	30	51	42.3 (39.2,45.4)	94.6 (93.2,96)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	399.5	22	890	30	50	42.3 (39.2,45.4)	94.7 (93.3,96.1)
Elbeltagi 2022[10]	HM-JACKarc (personal communication)	CRC	4	0	1	992	52	0.052419	401	0	940	52	0	0 (NE,NE)	100 (NE,NE)
Farrugia 2020[11]	HM JACKarc automated system	CRC	2	0	1	519	33	0.063584	10	28	395	5	91	84.8 (81.7,87.9)	81.3 (77.9,84.7)
Farrugia 2020[11]	HM JACKarc automated system	CRC	4	0	1	612	38	0.062092	10	33	472	5	102	86.8 (84.1,89.5)	82.2 (79.2,85.2)
Farrugia 2020[11]	HM JACKarc automated system	CRC	4	0	1	79	5	0.063291	10	5	68	0	6	100 (NE,NE)	91.9 (85.9,97.9)

Faux 2022[12]	HM-JACKarc analyser	CRC	4	0	2	175	6	0.034286	10	4	141	2	28	66.7 (59.7,73.7)	83.4 (77.9,88.9)
Gerrard 2023[13]	HM-JACKarc analyser	CRC	1	0	1	2260	69	0.030531	10	58	1696	11	495	84.1 (82.6,85.6)	77.4 (75.7,79.1)
Gerrard 2023[13]	HM-JACKarc analyser	CRC	1	0	1	3426	135	0.039405	10	126	2567	9	724	93.3 (92.5,94.1)	78 (76.6,79.4)
Gerrard 2023[13]	HM-JACKarc analyser	CRC	1	1 (unclear if IDA or anaemia)	1	567	38	0.067019	10	31	363	7	166	81.6 (78.4,84.8)	68.6 (64.8,72.4)
Gerrard 2023[13]	HM-JACKarc analyser	CRC	1	1 (unclear if IDA or anaemia)	1	480	29	0.060417	10	27	271	2	180	93.1 (90.8,95.4)	60.1 (55.7,64.5)
Godber 2016[14]	HM JACKarc analyser	CRC	4	0	1	484	11	0.022727	10	11	362	0	111	100 (NE,NE)	76.5 (72.7,80.3)
Johnstone 2022a[15]	HM-JACKarc confirmed by author	CRC	1	0	2	4737	61	0.012877	10	53	3763	5	916	91.8 (91,92.6)	80.4 (79.3,81.5)
Johnstone 2022a[15]	HM-JACKarc confirmed by author	CRC	1	1 - not anaemic	2	3238	32	0.009883	10	31	2631	1	606	96.9 (96.3,97.5)	81.3 (80,82.6)
Johnstone 2022a[15]	HM-JACKarc confirmed by author	CRC	1	1 - anaemia. anaemic (male < 130 mg/L, female < 120 mg/L) based on WHO guidelines	2	793	26	0.032787	10	22	559	4	208	84.6 (82.1,87.1)	72.9 (69.8,76)
Johnstone 2022a[15]	HM-JACKarc confirmed by author	CRC	1	0	2	4737	61	0.012877	150	42	4387	16	292	72.4 (71.1,73.7)	93.8 (93.1,94.5)

Johnstone 2022a[15]	HM-JACKarc confirmed by author	CRC	1	0	2	4737	61	0.012877	400	34	4492	24	187	58.6 (57.2,60)	96 (95.4,96.6)
MacDonald 2022[16]	HM-JACKarc	CRC	1	0	2	5250	151	0.028762	10	132	3399	19	1700	87.4 (86.5,88.3)	66.6 (65.3,67.9)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	2	102	2611	3	2665	97.1 (96.7,97.5)	49.5 (48.2,50.8)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	7	93	4004	12	1272	88.6 (87.8,89.4)	75.9 (74.8,77)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	10	91	4190	14	1086	86.7 (85.8,87.6)	79.4 (78.3,80.5)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	20	87	4459	18	817	82.9 (81.9,83.9)	84.5 (83.5,85.5)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	50	78	4706	27	570	74.3 (73.1,75.5)	89.2 (88.4,90)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	100	73	4846	32	430	69.5 (68.3,70.7)	91.8 (91.1,92.5)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	150	67	4924	38	352	63.8 (62.5,65.1)	93.3 (92.6,94)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	200	65	4963	40	313	61.9 (60.6,63.2)	94.1 (93.5,94.7)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	250	62	4990	43	286	59 (57.7,60.3)	94.6 (94,95.2)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	300	59	5015	46	261	56.2 (54.9,57.5)	95.1 (94.5,95.7)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	350	57	5029	48	247	54.3 (53,55.6)	95.3 (94.7,95.9)
Mowat 2021[17] & 2019[18]	HM JACKarc	CRC	1	0	2	5381	105	0.019513	400	56	5040	49	236	53.3 (52,54.6)	95.5 (94.9,96.1)
Nicholson 2019 [19]	HM JACKarc	CRC	4	0	2	238	7	0.029412	7	6	206	1	25	85.7 (81.3,90.1)	89.2 (85.3,93.1)
Nicholson 2019 [19]	HM JACKarc	CRC	4	0	2	238	7	0.029412	10	6	209	1	22	85.7 (81.3,90.1)	90.5 (86.8,94.2)

Nicholson 2019 [19]	HM JACKarc	CRC	4	0	2	238	7	0.029412	20	5	214	2	17	71.4 (65.7,77.1)	92.6 (89.3,95.9)
Nicholson 2019 [19]	HM JACKarc	CRC	4	0	2	238	7	0.029412	50	4	221	3	10	57.1 (50.8,63.4)	95.7 (93.1,98.3)
Nicholson (2020)	HM JACKarc	CRC	4	0	2	9896	105	0.01061	7	96	8792	9	999	91.4 (90.8,92)	89.8 (89.2,90.4)
Nicholson 2020[20]	HM JACKarc	CRC	4	2	2	4101	65	0.01585	7	60	3548	5	488	92.3 (91.5,93.1)	87.9 (86.9,88.9)
Nicholson 2020[20]	HM JACKarc	CRC	4	3	2	5795	40	0.006903	7	36	5243	4	512	90 (89.2,90.8)	91.1 (90.4,91.8)
Nicholson 2020[20]	HM JACKarc	CRC	4	0	2	9896	105	0.01061	10	95	8939	10	852	90.5 (89.9,91.1)	91.3 (90.7,91.9)
Nicholson 2020[20]	HM JACKarc	CRC	4	2	2	4101	65	0.01585	10	59	3624	6	412	90.8 (89.9,91.7)	89.8 (88.9,90.7)
Nicholson 2020[20]	HM JACKarc	CRC	4	3	2	5795	40	0.006903	10	36	5318	4	437	90 (89.2,90.8)	92.4 (91.7,93.1)
Nicholson 2020[20]	HM JACKarc	CRC	4	0	2	9896	105	0.01061	20	89	9174	16	617	84.8 (84.1,85.5)	93.7 (93.2,94.2)
Nicholson 2020[20]	HM JACKarc	CRC	4	2	2	4101	65	0.01585	20	54	3725	11	311	83.1 (82,84.2)	92.3 (91.5,93.1)
Nicholson 2020[20]	HM JACKarc	CRC	4	3	2	5795	40	0.006903	20	35	5444	5	311	87.5 (86.6,88.4)	94.6 (94,95.2)
Nicholson 2020[20]	HM JACKarc	CRC	4	0	2	9896	105	0.01061	50	78	9439	27	352	74.3 (73.4,75.2)	96.4 (96,96.8)
Nicholson 2020[20]	HM JACKarc	CRC	4	2	2	4101	65	0.01585	50	48	3854	17	182	73.8 (72.5,75.1)	95.5 (94.9,96.1)
Nicholson 2020[20]	HM JACKarc	CRC	4	3	2	5795	40	0.006903	50	30	5577	10	178	75 (73.9,76.1)	96.9 (96.5,97.3)
Nicholson 2020[20]	HM JACKarc	CRC	4	0	2	9896	105	0.01061	100	64	9556	41	235	61 (60,62)	97.6 (97.3,97.9)
Nicholson 2020[20]	HM JACKarc	CRC	4	2	2	4101	65	0.01585	100	39	3907	26	129	60 (58.5,61.5)	96.8 (96.3,97.3)

Nicholson 2020[20]	HM JACKarc	CRC	4	3	2	5795	40	0.006903	100	25	5646	15	109	62.5 (61.3,63.7)	98.1 (97.7,98.5)
Nicholson 2020[20]	HM JACKarc	CRC	4	0	2	9896	105	0.01061	120	60	9576	45	215	57.1 (56.1,58.1)	97.8 (97.5,98.1)
Nicholson 2020[20]	HM JACKarc	CRC	4	2	2	4101	65	0.01585	120	36	3923	29	113	55.4 (53.9,56.9)	97.2 (96.7,97.7)
Nicholson 2020[20]	HM JACKarc	CRC	4	3	2	5795	40	0.006903	120	24	5657	16	98	60 (58.7,61.3)	98.3 (98,98.6)
Nicholson 2020[20]	HM JACKarc	CRC	4	0	2	9896	105	0.01061	150	57	9605	48	186	54.3 (53.3,55.3)	98.1 (97.8,98.4)
Nicholson 2020[20]	HM JACKarc	CRC	4	2	2	4101	65	0.01585	150	33	3935	32	101	50.8 (49.3,52.3)	97.5 (97,98)
Nicholson 2020[20]	HM JACKarc	CRC	4	3	2	5795	40	0.006903	150	24	5669	16	86	60 (58.7,61.3)	98.5 (98.2,98.8)
Tang 2022[21]	HM-JACKarc system	CRC	4	0	1	603	20	0.033167	10	18	485	2	98	90 (87.6,92.4)	83.2 (80.2,86.2)
Tang 2022[21]	HM-JACKarc system	CRC	4	1 - IDA	1	78	1	0.012821	10	1	59	0	18	100 (NE,NE)	76.6 (67.2,86)
Turvill 2018[22]	HM-JACKarc- single FIT	CRC	2	0	unclear	505	27	0.053465	12	23	423	4	55	84.6 (81.5,87.7)	88.5 (85.7,91.3)
Turvill 2021[23]	HM JACKarc	CRC	4	0	unclear	5040	151	0.02996	2	140	2970	11	1919	92.7 (92,93.4)	60.7 (59.4,62)
Turvill 2021[23]	HM JACKarc	CRC	4	0	unclear	5040	151	0.02996	10	132	3956	19	933	87.4 (86.5,88.3)	80.9 (79.8,82)
Turvill 2021[23]	HM JACKarc	CRC	4	3	unclear	2798	62	0.022159	16	54	2342	8	394	87.1 (85.9,88.3)	85.6 (84.3,86.9)
Turvill 2021[23]	HM JACKarc	CRC	4	0	unclear	5040	151	0.02996	19	129	4165	22	724	85.4 (84.4,86.4)	85.2 (84.2,86.2)
Turvill 2021[23]	HM JACKarc	CRC	4	7 - â‰¥60 years	unclear	3823	121	0.031651	19	101	3162	20	540	83.5 (82.3,84.7)	85.4 (84.3,86.5)
Turvill 2021[23]	HM JACKarc	CRC	4	5 - Drug use	unclear	1356	51	0.037611	19	42	1051	9	254	82.4 (80.4,84.4)	80.5 (78.4,82.6)

				(antiplatelets, anticoagulants NSAIDs)											
Turvill 2021[23]	HM JACKarc	CRC	4	2	unclear	2242	89	0.039697	21	76	1802	13	351	85.4 (83.9,86.9)	83.7 (82.2,85.2)
Turvill 2021[23]	HM JACKarc	CRC	4	1 - IDA	unclear	559	34	0.060823	21	28	428	6	97	82.4 (79.2,85.6)	81.5 (78.3,84.7)
Turvill 2021[23]	HM JACKarc	CRC	4	0	unclear	5040	151	0.02996	30	121	4288	30	601	80.1 (79,81.2)	87.7 (86.8,88.6)
Turvill 2021[23]	HM JACKarc	CRC	4	1- non-IDA	unclear	544	25	0.045956	30	23	444	2	75	92 (89.7,94.3)	85.5 (82.5,88.5)
Turvill 2021[23]	HM JACKarc	CRC	4	7 - <60 years	unclear	1217	30	0.024651	37	27	1037	3	150	90 (88.3,91.7)	87.4 (85.5,89.3)
Turvill 2021[23]	HM JACKarc	CRC	4	0	unclear	5040	151	0.02996	100	100	4532	51	357	66.2 (64.9,67.5)	92.7 (92,93.4)
Turvill 2021[23]	HM JACKarc	CRC	4	0	unclear	5040	151	0.02996	300	80	4649	71	240	53 (51.6,54.4)	95.1 (94.5,95.7)
Withrow 2022[24]	HM JACKarc	CRC	3	0	2	11142	89	0.007988	2	85	9209	4	1844	95.5 (95.1,95.9)	83.3 (82.6,84)
Withrow 2022[24]	HM JACKarc	CRC	4	0	2	16604	139	0.008371	2	134	13752	5	2713	96.4 (96.1,96.7)	83.5 (82.9,84.1)
Withrow 2022[24]	HM JACKarc	CRC	4	2	2	7019	83	0.011825	2	80	5678	3	1258	96.4 (96,96.8)	81.9 (81,82.8)
Withrow 2022[24]	HM JACKarc	CRC	4	3	2	9585	57	0.005947	2	54	8074	2	1455	96.4 (96,96.8)	84.7 (84,85.4)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age <40 years	2	1390	9	0.006475	2	9	1231	0	150	100 (NE,NE)	89.1 (87.5,90.7)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >40 years	2	15214	130	0.008545	2	125	12521	5	2563	96.2 (95.9,96.5)	83 (82.4,83.6)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >50 years	2	12936	118	0.009122	2	113	10483	5	2335	95.8 (95.5,96.1)	81.8 (81.1,82.5)

Withrow 2022[24]	HM JACKarc	CRC	4	7 age >60 years	2	8755	98	0.011194	2	93	6823	5	1834	94.9 (94.4,95.4)	78.8 (77.9,79.7)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >70 years	2	3043	77	0.025304	2	73	1536	4	1430	94.8 (94,95.6)	51.8 (50,53.6)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >80 years	2	2527	41	0.016225	2	39	1701	2	785	95.1 (94.3,95.9)	68.4 (66.6,70.2)
Withrow 2022[24]	HM JACKarc	CRC	3	0	2	11142	89	0.007988	10	82	10142	7	911	92.1 (91.6,92.6)	91.8 (91.3,92.3)
Withrow 2022[24]	HM JACKarc	CRC	4	0	2	16604	139	0.008371	10	128	15064	11	1401	92.1 (91.7,92.5)	91.5 (91.1,91.9)
Withrow 2022[24]	HM JACKarc	CRC	4	2	2	7019	83	0.011825	10	77	6262	6	674	92.8 (92.2,93.4)	90.3 (89.6,91)
Withrow 2022[24]	HM JACKarc	CRC	4	3	2	9585	57	0.005947	10	51	8802	5	727	91.1 (90.5,91.7)	92.4 (91.9,92.9)
Withrow 2022[24]	HM JACKarc	CRC	4	1 - Low Haemoglobin (<130 g/L in men, <120g/L in women)	2	5076	72	0.014184	10	69	4404	3	600	95.8 (95.2,96.4)	88 (87.1,88.9)
Withrow 2022[24]	HM JACKarc	CRC	4	1 & 2	2	2091	46	0.021999	10	43	1749	3	296	93.5 (92.4,94.6)	85.5 (84,87)
Withrow 2022[24]	HM JACKarc	CRC	4	1 & 2	2	1141	36	0.031551	10	33	914	3	191	91.7 (90.1,93.3)	82.7 (80.5,84.9)
Withrow 2022[24]	HM JACKarc	CRC	4	1 & 2	2	494	23	0.046559	10	22	372	1	99	95.7 (93.9,97.5)	79 (75.4,82.6)
Withrow 2022[24]	HM JACKarc	CRC	4	1 & 2	2	216	14	0.064815	10	14	146	0	56	100 (NE,NE)	72.3 (66.3,78.3)
Withrow 2022[24]	HM JACKarc	CRC	4	1 & 2	2	89	9	0.101124	10	9	57	0	23	100 (NE,NE)	71.2 (61.8,80.6)

Withrow 2022[24]	HM JACKarc	CRC	4	1 &3	2	2758	25	0.009065	10	25	2444	0	289	100 (NE,NE)	89.4 (88.3,90.5)
Withrow 2022[24]	HM JACKarc	CRC	4	1 &3	2	1297	13	0.010023	10	13	1130	0	154	100 (NE,NE)	88 (86.2,89.8)
Withrow 2022[24]	HM JACKarc	CRC	4	1 &3	2	491	6	0.01222	10	6	410	0	75	100 (NE,NE)	84.5 (81.3,87.7)
Withrow 2022[24]	HM JACKarc	CRC	4	1 &3	2	189	3	0.015873	10	3	148	0	38	100 (NE,NE)	79.6 (73.9,85.3)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age <40 years	2	1390	9	0.006475	10	8	1290	1	91	88.9 (87.2,90.6)	93.4 (92.1,94.7)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >40 years	2	15214	130	0.008545	10	120	13774	10	1310	92.3 (91.9,92.7)	91.3 (90.9,91.7)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >50 years	2	12936	118	0.009122	10	108	11629	10	1189	91.5 (91,92)	90.7 (90.2,91.2)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >60 years	2	8755	98	0.011194	10	88	7705	10	952	89.8 (89.2,90.4)	89 (88.3,89.7)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >70 years	2	5863	77	0.013133	10	69	5038	8	748	89.6 (88.8,90.4)	87.1 (86.2,88)
Withrow 2022[24]	HM JACKarc	CRC	4	7 age >80 years	2	2533	41	0.016186	10	36	2071	5	421	87.8 (86.5,89.1)	83.1 (81.6,84.6)
OC-Sensor															
Archer 2022[25]	OC-Sensor in other Sheffield article, Ball 2022 - but NR for this study	CRC	4	0	1	166	11	0.066265	10	10	52	1	103	90.9 (86.5,95.3)	33.6 (26.4,40.8)
Archer 2022[25]	OC-Sensor in other Sheffield article, Ball 2022 - but NR for this study	CRC	4	0	1	166	11	0.066265	60	6	115	5	40	54.6 (47,62.2)	74.2 (67.5,80.9)
Archer 2022[25]	OC-Sensor in other Sheffield article, Ball 2022 - but NR for this study	CRC	4	0	1	166	11	0.066265	100	6	122	5	33	54.6 (47,62.2)	78.7 (72.5,84.9)
Ayling 2019[26]	OC-Sensor (analyser NR)	CRC	4	1 - low haemoglobin (not defined)	1	178	7	0.039326	10	5	164	2	7	71.4 (64.8,78)	95.9 (93,98.8)

Ayling 2019[26]	OC-Sensor (analyser NR)	CRC	4	1 - IDA	1	137	6	0.043796	10	4	125	2	6	68.7 (60.9,76.5)	95.4 (91.9,98.9)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	0	2	3506	45	0.012835	10	41	2794	4	667	91.1 (90.2,92)	80.7 (79.4,82)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	2	2	1566	25	0.015964	10	21	1220	4	321	84 (82.2,85.8)	79.2 (77.2,81.2)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	3	2	1940	20	0.010309	10	20	1574	0	346	100 (NE,NE)	82 (80.3,83.7)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	0	2	3506	45	0.012835	20	39	3021	6	440	86.7 (85.6,87.8)	87.3 (86.2,88.4)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	2	2	1566	25	0.015964	20	20	1008	5	533	80 (78,82)	65.4 (63,67.8)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	3	2	1940	20	0.010309	20	19	1705	1	215	95 (94,96)	88.8 (87.4,90.2)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	0	2	3506	45	0.012835	50	33	3217	12	244	73.3 (71.8,74.8)	93 (92.2,93.8)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	2	2	1566	25	0.015964	50	17	1412	8	129	68 (65.7,70.3)	91.6 (90.2,93)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	3	2	1940	20	0.010309	50	16	1807	4	113	80 (78.2,81.8)	94.1 (93.1,95.1)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	0	2	3506	45	0.012835	80	30	3287	15	174	66.7 (65.1,68.3)	95 (94.3,95.7)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	2	2	1566	25	0.015964	80	16	1447	9	94	64 (61.6,66.4)	93.9 (92.7,95.1)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	3	2	1940	20	0.010309	80	14	1839	6	81	70 (68,72)	95.8 (94.9,96.7)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	0	2	3506	45	0.012835	100	30	3315	15	146	66.7 (65.1,68.3)	95.8 (95.1,96.5)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	2	2	1566	25	0.015964	100	16	1458	9	83	64 (61.6,66.4)	94.6 (93.5,95.7)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	3	2	1940	20	0.010309	100	14	1857	6	63	70 (68,72)	96.7 (95.9,97.5)

Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	0	2	3506	45	0.012835	120	28	3330	17	131	62.2 (60.6,63.8)	96.2 (95.6,96.8)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	2	2	1566	25	0.015964	120	15	1467	10	74	60 (57.6,62.4)	95.2 (94.1,96.3)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	3	2	1940	20	0.010309	120	13	1862	7	58	65 (62.9,67.1)	97 (96.2,97.8)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	0	2	3506	45	0.012835	150	24	3461	21	107	53.3 (51.6,55)	97 (96.4,97.6)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	2	2	1566	25	0.015964	150	13	1486	12	55	52 (49.5,54.5)	96.4 (95.5,97.3)
Ball 2022[27]	OC-Sensor Pledia analyser	CRC	4	3	2	1940	20	0.010309	150	11	1868	9	52	55 (52.8,57.2)	97.3 (96.6,98)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	0	2	2892	17	0.005878	10	16	2458	1	417	94.1 (93.2,95)	85.5 (84.2,86.8)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	2	2	1286	11	0.008554	10	10	1074	1	201	90.9 (89.3,92.5)	84.2 (82.2,86.2)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	3	2	1606	6	0.003736	10	6	1384	0	216	100 (NE,NE)	86.5 (84.8,88.2)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	0	2	2892	17	0.005878	20	15	2623	2	252	88.2 (87,89.4)	91.2 (90.2,92.2)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	2	2	1286	11	0.008554	20	9	1139	2	136	81.8 (79.7,83.9)	89.3 (87.6,91)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	3	2	1606	6	0.003736	20	6	1485	0	115	100 (NE,NE)	92.8 (91.5,94.1)

Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	0	2	2892	17	0.005878	50	11	2736	6	139	64.7 (63,66.4)	95.2 (94.4,96)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	2	2	1286	11	0.008554	50	7	1196	4	79	63.6 (61,66.2)	93.8 (92.5,95.1)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	3	2	1606	6	0.003736	50	4	1541	2	59	66.7 (64.4,69)	96.3 (95.4,97.2)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	0	2	2892	17	0.005878	80	10	2778	7	97	58.8 (57,60.6)	96.6 (95.9,97.3)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	2	2	1286	11	0.008554	80	7	1218	4	57	63.6 (61,66.2)	95.5 (94.4,96.6)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	3	2	1606	6	0.003736	80	3	1560	3	40	50 (47.6,52.4)	97.5 (96.7,98.3)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	0	2	2892	17	0.005878	100	10	2797	7	78	58.8 (57,60.6)	97.3 (96.7,97.9)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	2	2	1286	11	0.008554	100	7	1227	4	48	63.6 (61,66.2)	96.2 (95.2,97.2)
Ball 2022[27] (personal communicaton)	OC-Sensor Pledia analyser	CRC	3	3	2	1606	6	0.003736	100	3	1571	3	29	50 (47.6,52.4)	98.2 (97.5,98.9)
Ball 2022[27] (personal	OC-Sensor Pledia analyser	CRC	3	0	2	2892	17	0.005878	120	10	2807	7	68	58.8 (57,60.6)	97.6 (97,98.2)

commun icaton)															
Ball 2022[27] (personal commun icaton)	OC-Sensor Pledia analyser	CRC	3	2	2	1286	11	0.008554	120	7	1233	4	42	63.6 (61,66.2)	96.7 (95.7,97.7)
Ball 2022[27] (personal commun icaton)	OC-Sensor Pledia analyser	CRC	3	3	2	1606	6	0.003736	120	3	1574	3	26	50 (47.6,52.4)	98.4 (97.8,99)
Ball 2022[27] (personal commun icaton)	OC-Sensor Pledia analyser	CRC	3	0	2	2892	17	0.005878	150	9	2827	8	48	52.9 (51.1,54.7)	98.3 (97.8,98.8)
Ball 2022[27] (personal commun icaton)	OC-Sensor Pledia analyser	CRC	3	2	2	1286	11	0.008554	150	6	1247	5	28	54.5 (51.8,57.2)	97.8 (97,98.6)
Ball 2022[27] (personal commun icaton)	OC-Sensor Pledia analyser	CRC	3	3	2	1606	6	0.003736	150	3	1581	3	19	50 (47.6,52.4)	98.8 (98.3,99.3)
Benton 2022[1]	OC Sensor PLEDIA	CRC	2	0	1	233	7	0.030043	1	5	103	2	123	71.4 (65.6,77.2)	45.6 (39.2,52)
Benton 2022[1]	OC Sensor PLEDIA	CRC	2	0	1	233	7	0.030043	10	5	194	2	32	71.4 (65.6,77.2)	85.8 (81.3,90.3)
Benton 2022[1]	OC Sensor PLEDIA	CRC	2	0	1	233	7	0.030043	100	4	218	3	8	57.1 (50.7,63.5)	96.5 (94.1,98.9)
Bujanda 2018[28]	OC-Sensor (analyser NR)	CRC	4	5 - with aspirin	1	485	51	0.105155	20	45	291	6	143	88 (85.1,90.9)	67 (62.8,71.2)
Bujanda 2018[28]	OC-Sensor (analyser NR)	CRC	4	5 - without aspirin	1	2567	299	0.116478	20	275	1610	24	658	92 (91,93)	71 (69.2,72.8)
Cama 2022[29]	OC-Sensor iO	CRC	1	0	2	5341	74	0.013855	4	69	2950	5	2317	93 (92.3,93.7)	56 (54.7,57.3)

Cama 2022[29]	OC-Sensor iO	CRC	1	0	2	5341	74	0.013855	10	67	4108	7	1159	90.5 (89.7,91.3)	78 (76.9,79.1)
Cama 2022[29]	OC-Sensor iO	CRC	1	0	2	5341	74	0.013855	100	53	5004	21	263	71.6 (70.4,72.8)	95 (94.4,95.6)
Chapman 2021[4]	OC-Sensor DIANA	CRC	4	0	1	732	38	0.051913	4	37	444	1	250	97 (95.8,98.2)	64 (60.5,67.5)
Chapman 2021[4]	OC-Sensor DIANA	CRC	4	0	1	732	38	0.051913	10	34	514	4	180	89 (86.7,91.3)	74 (70.8,77.2)
Chapman 2021[4]	OC-Sensor DIANA	CRC	4	0	1	732	38	0.051913	18.2	33	548	5	146	87 (84.6,89.4)	79 (76,82)
Chapman 2021[4]	OC-Sensor DIANA	CRC	4	0	1	732	38	0.051913	150	24	652	14	42	63 (59.5,66.5)	94 (92.3,95.7)
Crooks 2023[30]	OC-Sensor iO (Bailey 2021a)	CRC	1	0	2	33694	514	0.015255	4	488	20461	26	12719	94.9 (94.7,95.1)	61.7 (61.2,62.2)
Crooks 2023[30]	OC-Sensor iO (Bailey 2021a)	CRC	1	0	2	33694	514	0.015255	10	461	26004	53	7176	89.7 (89.4,90)	78.4 (78,78.8)
Crooks 2023[30]	OC-Sensor iO (Bailey 2021a)	CRC	1	0	2	33694	514	0.015255	20	437	28353	77	4827	85 (84.6,85.4)	85.5 (85.1,85.9)
Crooks 2023[30]	OC-Sensor iO (Bailey 2021a)	CRC	1	0	2	33694	514	0.015255	40	396	30061	118	3119	77 (76.6,77.4)	90.6 (90.3,90.9)
Crooks 2023[30]	OC-Sensor iO (Bailey 2021a)	CRC	1	0	2	33694	514	0.015255	100	329	31552	185	1628	64 (63.5,64.5)	95.1 (94.9,95.3)
Georgiou Delisle 2022[31]	OC Sensor iO	CRC	1	0	2	4187	61	0.014569	4	59	1622	2	2504	96.7 (96.2,97.2)	39.3 (37.8,40.8)
Georgiou Delisle 2022[31]	OC Sensor iO	CRC	1	0	2	4187	61	0.014569	9.5	58	3099	3	1027	95.1 (94.4,95.8)	75.1 (73.8,76.4)
Georgiou Delisle 2022[31]	OC Sensor iO	CRC	1	0	2	4187	61	0.014569	150	40	3874	21	252	65.6 (64.2,67)	93.9 (93.2,94.6)
Juul 2018[32]	OC-Sensor DIANA	CRC	4	0	2	3462	54	0.015598	10	51	2919	3	489	94.4 (93.6,95.2)	85.7 (84.5,86.9)
Juul 2018[32]	OC-Sensor DIANA	CRC	4	1	2	424	49	0.115566	10	10	298	39	77	20.4 (16.6,24.2)	79.5 (75.7,83.3)

Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	4	79	2558	11	948	87.8 (86.7,88.9)	73 (71.5,74.5)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	6	78	2664	12	842	86.7 (85.6,87.8)	76.1 (74.7,77.5)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	10	75	2807	15	699	83.3 (82.1,84.5)	80.1 (78.8,81.4)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	20	73	2997	17	509	81.1 (79.8,82.4)	85.5 (84.3,86.7)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	50	67	3209	23	297	74.4 (73,75.8)	91.6 (90.7,92.5)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	80	61	3269	29	237	67.8 (66.3,69.3)	93.3 (92.5,94.1)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	100	58	3298	32	208	64.4 (62.8,66)	94.1 (93.3,94.9)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	120	55	3319	35	187	61.1 (59.5,62.7)	94.7 (94,95.4)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	150	52	3335	38	171	57.8 (56.2,59.4)	95.2 (94.5,95.9)
Laszlo 2021[33]	OC-Sensor iO	CRC	4	0	2	3596	90	0.025028	200	49	3352	41	154	54.4 (52.8,56)	95.7 (95,96.4)
Maclean 2021a[34]	OC-Sensor PLEDIA	CRC	4	0	2	358	12	0.03352	10	12	246	0	100	100 (NE,NE)	71.1 (66.4,75.8)
Maclean 2021a[34]	OC-Sensor PLEDIA	CRC	4	0	2	358	12	0.03352	150	5	318	7	28	41.7 (36.6,46.8)	91.9 (89.1,94.7)
Morales-Arraez 2018[35]	OC-Sensor	CRC	4	1 - Hb<11.9 g/dL in men and Hb<10.9 g/dL in women, and ferritin	1	245	28	0.114286	10	26	124	2	93	92.9 (89.7,96.1)	57.1 (50.9,63.3)

				â‰¤ 30 g/dL											
Mowat 2016[36]	OC-Sensor io	CRC	4	0	1	750	28	0.037333	4	28	313	0	409	100 (NE,NE)	43.4 (39.9,46.9)
Mowat 2016[36]	OC-Sensor io	CRC	4	0	1	750	28	0.037333	10	25	571	3	151	89.3 (87.1,91.5)	79.1 (76.2,82)
Pin Vieto 2021[37]	OC-Sensor analyser NR	CRC	4	0	2	4543	73	0.016069	10	59	3728	14	742	80.6 (79.5,81.7)	83.4 (82.3,84.5)
Pin Vieto 2021[37]	OC-Sensor analyser NR	CRC	4	0	2	4543	73	0.016069	20	57	3916	16	554	77.8 (76.6,79)	87.6 (86.6,88.6)
Rodriguez-Alonso 2018[38]	OC-Sensor MICRO	CRC	4	5 - PPI users	1	525	15	0.028571	20	14	434	1	76	93.3 (91.2,95.4)	85.1 (82.1,88.1)
Rodriguez-Alonso 2018[38]	OC-Sensor MICRO	CRC	4	5 - PPI non-users	1	477	15	0.031447	20	14	404	1	58	93.3 (91.1,95.5)	87.4 (84.4,90.4)
Rodriguez-Alonso 2020[39]	OC-Sensor MICRO desktop analyser	CRC	4 - referred to colonoscopy in Spain	1 iron deficiency anaemia	1	120	9	0.075	15	9	86	0	25	100 (NE,NE)	77.5 (70,85)
QuikRead go															
Maclean 2021b[40]	QuikRead go	CRC	2	0	2	553	14	0.025316	10	13	378	1	161	92.9 (90.8,95)	70.1 (66.3,73.9)
Maclean 2021b[40]	QuikRead go	CRC	2	0	2	553	14	0.025316	100	10	510	4	29	71.4 (67.6,75.2)	94.6 (92.7,96.5)
Maclean 2021b[40]	QuikRead go	CRC	2	0	2	553	14	0.025316	150	8	517	6	22	57.1 (53,61.2)	95.9 (94.2,97.6)
Tsapournas 2020[41]	QuikRead go (inc referrals from primary and secondary care)	CRC	4	0	1	242	13	0.053719	10	12	177	1	52	92.3 (88.9,95.7)	77.3 (72,82.6)
Tsapournas 2020[41]	QuikRead go (inc referrals from primary and secondary care)	CRC	4	0	1	242	13	0.053719	15	12	187	1	42	92.3 (88.9,95.7)	81.7 (76.8,86.6)
Tsapournas 2020[41]	QuikRead go (inc referrals from primary and secondary care)	CRC	4	0	1	242	13	0.053719	20	11	198	2	31	84.6 (80.1,89.1)	86.5 (82.2,90.8)

IDK															
Sieg 1999[42]	IDK	CRC	4	0	1	621	23	0.037037	2	19	483	4	115	82.6 (79.6,85.6)	80.8 (77.7,83.9)
Sieg 1999[42]	IDK	CRC	4	0	1	621	23	0.037037	2	20	527	3	71	87 (84.4,89.6)	88.1 (85.6,90.6)
NS-Prime															
Benton 2022[1]	NS-Prime	CRC	2	0	1	233	7	0.030043	3	6	72	1	154	85.7 (81.2,90.2)	31.9 (25.9,37.9)
Benton 2022[1]	NS-Prime	CRC	2	0	1	233	7	0.030043	10	5	189	2	37	71.4 (65.6,77.2)	83.6 (78.8,88.4)
Benton 2022[1]	NS-Prime	CRC	2	0	1	233	7	0.030043	100	4	220	3	6	57.1 (50.7,63.5)	97.3 (95.2,99.4)
Dual FIT, all tests															
Gerrard 2023[13]	DUAL FIT HM-JACKarc (either positive)	CRC	1	0	1	2637	88	0.033371	10	85	1815	3	734	96.6 (95.9,97.3)	71.2 (69.5,72.9)
Hunt 2022[43]	DUAL FIT OC-Sensor (both positive) analyser NR	CRC	4	0	2	28622	317	0.011075	10	290	23097	27	5208	91.5 (91.2,91.8)	81.6 (81.2,82)
Hunt 2022[43]	DUAL FIT OC-Sensor (either positive) analyser NR	CRC	4	0	2	28622	317	0.011075	10	310	18738	7	9567	97.8 (97.6,98)	66.2 (65.7,66.7)
Tsapournas 2020[41]	DUAL FIT QuikRead go (either positive) (inc referrals from primary and secondary care)	CRC	4	0	1	242	13	0.053719	10	13	164	0	65	100 (NE,NE)	71.4 (65.7,77.1)
Tsapournas 2020[41]	DUAL FIT QuikRead go (either positive) (inc referrals from primary and secondary care)	CRC	4	0	1	242	13	0.053719	15	12	176	1	53	92.3 (88.9,95.7)	76.8 (71.5,82.1)
Tsapournas 2020[41]	DUAL FIT QuikRead go (either positive) (inc referrals from primary and secondary care)	CRC	4	0	1	242	13	0.053719	20	12	187	1	42	92.3 (88.9,95.7)	81.7 (76.8,86.6)
Turvill 2018[22]	dual FIT (both positive) HM-JACKarc	CRC	2	0	unclear	476	27	0.056723	2	25	383	2	66	91.7 (89.2,94.2)	85.2 (82,88.4)

Turvill 2018[22]	dual FIT (either positive) HM-JACKarc	CRC	2	0	unclear	476	27	0.056723	43	24	407	3	42	87.5 (84.5,90.5)	90.7 (88.1,93.3)

CRC, Colorectal cancer; FN, false negative; FP, False positive; No., number; Pop., population; Prev, prevalence; Pts, patients; Ref Stand, reference standard; TN, true negative; TP, true positive

Table 2 Data entering the statistical syntheses of the diagnostic test accuracy of FITs for detection of advanced adenomas and inflammatory bowel disease

Author, year	Test	Out-come	Pop. type	Sub-group	Ref Stand	No. Pts	CRC cases	Prev CRC	Thres-hold (µg/g)	TP	TN	FN	FP	Sensitivity	Specificity
Advanced Adenoma															
Sieg 1999[42]	IDK	AA	4	0	1	621	37	0.059581	2	27	477	10	107	73 (69.5,76.5)	81.7 (78.7,84.7)
Sieg 1999[42]	IDK	AA	4	0	1	621	37	0.059581	2	20	513	17	71	54.1 (50.2,58)	87.8 (85.2,90.4)
D'Souza 2020a[6]	HM JACKarc analytical system	AA	1	0	1	298	4	0.013423	2	2	216	2	78	50 (44.3,55.7)	73.5 (68.5,78.5)
D'Souza 2020a[6]	HM JACKarc analytical system	AA	1	0	1	298	4	0.013423	10	2	265	2	29	50 (44.3,55.7)	90.1 (86.7,93.5)
D'Souza 2020a[6]	HM JACKarc analytical system	AA	2	0	1	160	4	0.025	2	2	105	2	51	40 (32.4,47.6)	67.4 (60.1,74.7)
D'Souza 2020a[6]	HM JACKarc analytical system	AA	2	0	1	160	4	0.025	10	2	139	2	17	40 (32.4,47.6)	89.4 (84.6,94.2)
D'Souza 2020a[6]	HM JACKarc analytical system	AA	3	0	1	138	0	0	2	0	111	0	27	50 (41.7,58.3)	80.2 (73.6,86.8)
D'Souza 2020a[6]	HM JACKarc analytical system	AA	3	0	1	138	0	0	10	0	126	0	12	50 (41.7,58.3)	91.2 (86.5,95.9)
D'Souza 2021c[8]	HM JACKarc analytical system	AA	4	0	1	9822	421	0.042863	2	277	6026	144	3375	65.8 (64.9,66.7)	64.1 (63.2,65)
D'Souza 2021c[8]	HM JACKarc analytical system	AA	4	0	1	9822	421	0.042863	10	191	7728	230	1673	45.4 (44.4,46.4)	82.2 (81.4,83)
D'Souza 2021c[8]	HM JACKarc analytical system	AA	4	0	1	9822	421	0.042863	150	93	8743	328	658	22.1 (21.3,22.9)	93 (92.5,93.5)
Gerrard 2023[13]	DUAL FIT HM-JACKarc (either positive)	AA	1	0	1	2637	97	0.036784	10	66	1788	31	752	68 (66.2,69.8)	70.4 (68.7,72.1)

Gerrard 2023[13]	HM-JACKarc analyser	AA	1	0	1	2260	105	0.04646	10	54	1655	51	500	51.4 (49.3,53.5)	76.8 (75.1,78.5)
Gerrard 2023[13]	HM-JACKarc analyser	AA	1	0	1	3426	136	0.039696	10	74	2514	62	776	54.4 (52.7,56.1)	76.4 (75,77.8)
Juul 2018[32]	OC-Sensor DIANA	AA	4	0	2	3462	68	0.019642	10	62	2916	6	478	91.2 (90.3,92.1)	85.9 (84.7,87.1)
MacDonald 2022[16]	HM-JACKarc	AA	1	0	2	5250	47	0.008952	10	31	3402	16	1801	63.8 (62.5,65.1)	65.4 (64.1,66.7)
Maclean 2021b[40]	QuikRead go	AA	2	0	2	553	29	0.052441	10	19	369	10	155	65.5 (61.5,69.5)	70.4 (66.6,74.2)
Maclean 2021b[40]	QuikRead go	AA	2	0	2	553	29	0.052441	100	6	491	23	33	20.7 (17.3,24.1)	93.7 (91.7,95.7)
Maclean 2021b[40]	QuikRead go	AA	2	0	2	553	29	0.052441	150	4	498	25	26	13.8 (10.9,16.7)	95 (93.2,96.8)
Mowat 2016[36]	OC-Sensor io	AA	4	0	1	750	40	0.053333	4	33	306	7	404	82.5 (79.8,85.2)	43.1 (39.6,46.6)
Mowat 2016[36]	OC-Sensor io	AA	4	0	1	750	40	0.053333	10	20	554	20	156	50 (46.4,53.6)	78 (75,81)
Mowat 2021[17] & 2019[18]	HM JACKarc	AA	4	0	1	1447	133	0.091914	10	102	636	31	678	76.7 (74.5,78.9)	48.4 (45.8,51)
Inflammatory bowel disease															
Sieg 1999[42]	IDK	IBD	4	0	1	621	22	0.035427	2	19	516	3	83	86.4 (83.7,89.1)	86.1 (83.4,88.8)
Sieg 1999[42]	IDK	IBD	4	0	1	621	22	0.035427	2	16	524	6	75	72.7 (69.2,76.2)	87.5 (84.9,90.1)
D'Souza 2020a[6]	HM JACKarc analytical system	IBD	1	0	1	298	12	0.040268	2	9	215	3	71	75 (70.1,79.9)	75.2 (70.3,80.1)
D'Souza 2020a[6]	HM JACKarc analytical system	IBD	1	0	1	298	12	0.040268	10	8	249	4	37	66.7 (61.3,72.1)	87.1 (83.3,90.9)
D'Souza 2020a[6]	HM JACKarc analytical system	IBD	2	0	1	160	9	0.05625	2	8	107	1	44	88.9 (84,93.8)	70.9 (63.9,77.9)

D'Souza 2020a[6]	HM JACKarc analytical system	IBD	2	0	1	160	9	0.05625	10	7	127	2	24	77.8 (71.4,84.2)	84.1 (78.4,89.8)
D'Souza 2020a[6]	HM JACKarc analytical system	IBD	3	0	1	138	3	0.021739	2	1	108	2	27	33.3 (25.4,41.2)	80 (73.3,86.7)
D'Souza 2020a[6]	HM JACKarc analytical system	IBD	3	0	1	138	3	0.021739	10	1	122	2	13	33.3 (25.4,41.2)	90.3 (85.4,95.2)
D'Souza 2021c[8]	HM JACKarc analytical system	IBD	4	0	1	9822	427	0.043474	2	312	6050	115	3345	73.1 (72.2,74)	64.4 (63.5,65.3)
D'Souza 2021c[8]	HM JACKarc analytical system	IBD	4	0	1	9822	427	0.043474	10	247	7779	180	1616	57.8 (56.8,58.8)	82.8 (82.1,83.5)
D'Souza 2021c[8]	HM JACKarc analytical system	IBD	4	0	1	9822	427	0.043474	150	157	8803	270	592	36.8 (35.8,37.8)	93.7 (93.2,94.2)
Gerrard 2023[13]	DUAL FIT HM-JACKarc (either positive)	IBD	1	0	1	2637	33	0.012514	10	30	1815	3	789	90.9 (89.8,92)	69.7 (67.9,71.5)
Gerrard 2023[13]	HM-JACKarc analyser	IBD	1	0	1	2260	59	0.026106	10	45	1693	14	508	76.3 (74.5,78.1)	76.9 (75.2,78.6)
Gerrard 2023[13]	HM-JACKarc analyser	IBD	1	0	1	3426	55	0.016054	10	50	2572	5	799	90.9 (89.9,91.9)	76.3 (74.9,77.7)
Juul 2018[32]	OC-Sensor DIANA	IBD	4	0	2	3462	31	0.008954	10	11	2902	20	529	35.5 (33.9,37.1)	84.6 (83.4,85.8)
MacDonald 2022[16]	HM-JACKarc	IBD	1	0	2	5250	131	0.024952	10	91	3378	40	1741	69.5 (68.3,70.7)	66 (64.7,67.3)
Maclean 2021b[40]	QuikRead go	IBD	2	0	2	553	9	0.016275	10	8	378	1	166	88.9 (86.3,91.5)	69.5 (65.7,73.3)
Maclean 2021b[40]	QuikRead go	IBD	2	0	2	553	9	0.016275	100	4	509	5	35	44.4 (40.3,48.5)	93.6 (91.6,95.6)
Maclean 2021b[40]	QuikRead go	IBD	2	0	2	553	9	0.016275	150	3	517	6	27	33.3 (29.4,37.2)	95 (93.2,96.8)
Mowat 2016[36]	OC-Sensor io	IBD	4	0	1	750	34	0.045333	4	29	308	5	408	85.3 (82.8,87.8)	43 (39.5,46.5)
Mowat 2016[36]	OC-Sensor io	IBD	4	0	1	750	34	0.045333	10	25	565	9	151	73.5 (70.3,76.7)	78.9 (76,81.8)

Mowat 2021[17] & 2019[18]	HM JACKarc	IBD	4	0	1	1447	68	0.046994	10	64	663	4	716	94.1 (92.9,95.3)	48.1 (45.5,50.7)

CRC, Colorectal cancer; FN, false negative; FP, False positive; No., number; Pop., population; Prev, prevalence; Pts, patients; Ref Stand, reference standard; TN, true negative; TP, true positive

1. Benton SC, Piggott C, Zahoor Z, O'Driscoll S, Fraser CG, D'Souza N*, et al.* A comparison of the faecal haemoglobin concentrations and diagnostic accuracy in patients suspected with colorectal cancer and serious bowel disease as reported on four different faecal immunochemical test systems. *Clinical Chemistry & Laboratory Medicine* 2022;60:1278-86.
2. MacLean W, Zahoor Z, O'Driscoll S, Piggott C, Whyte MB, Rockall T*, et al.* Comparison of the QuikRead gopoint-of-care faecal immunochemical test for haemoglobin with the FOB Gold Widelaboratory analyser to diagnose colorectal cancer in symptomatic patients. *Clinical Chemistry and Laboratory Medicine* 2022a;60(1):101-8.
3. Schwettmann L, Lied A, Eriksen R. Evaluation of the Sentinel-FOB gold faecal immunochemical test for the presence of haemoglobin using the automated Roche Cobas 8000 system. *Practical Laboratory Medicine* 2022;29:e00263.
4. Chapman CJ, Banerjea A, Humes DJ, Allen J, Oliver S, Ford A*, et al.* Choice of faecal immunochemical test matters: comparison of OC-Sensor and HM-JACKarc, in the assessment of patients at high risk of colorectal cancer. *Clinical Chemistry & Laboratory Medicine* 2021;59:721-8.
5. Cunin L, Khan AA, Ibrahim M, Lango A, Klimovskij M, Harshen R. FIT negative cancers: A right-sided problem? Implications for screening and whether iron deficiency anaemia has a role to play. *The Surgeon* 2021;19:27-32.
6. D'Souza N, Hicks G, Benton SC, Abulafi M. The diagnostic accuracy of the faecal immunochemical test for colorectal cancer in risk-stratified symptomatic patients. *Annals of the Royal College of Surgeons of England* 2020a;102:174-9.
7. D'Souza N, Delisle TG, Chen M, Benton SC, Abulafi M, the NFITSC. Faecal immunochemical testing in symptomatic patients to prioritize investigation: diagnostic accuracy from NICE FIT Study. *British Journal of Surgery* 2021a;108:804-10.
8. D'Souza N, Delisle TG, Chen M, Benton S, Abulafi M. Faecal immunochemical test is superior to symptoms in predicting pathology in patients with suspected colorectal cancer symptoms referred on a 2WW pathway: a diagnostic accuracy study. *Gut* 2021c;70:1130-8.
9. D'Souza N, Monahan K, Benton SC, Wilde L, Abulafi M, Group NFS*, et al.* Finding the needle in the haystack: the diagnostic accuracy of the faecal immunochemical test for colorectal cancer in younger symptomatic patients. *Colorectal Disease* 2021b;23:2539-49.
10. Elbeltagi A, Salama M, Boxall P, Roos J, Lim M. The Yield of Faecal Immunochemical Test in the Detection of Colorectal Cancer within a Fast-track Pathway at York, United Kingdom. *Turkish Journal of Colorectal Disease* 2022;32(3):178-85.
11. Farrugia A, Widlak M, Evans C, Smith SC, Arasaradnam R. Faecal immunochemical testing (FIT) in symptomatic patients: what are we missing? *Frontline Gastroenterol* 2020;11:28-33.
12. Faux JW, Cock K, Bromley R, Feldman M. Colorectal two-week wait service and quantitative FIT: it's not just about colon cancer. *Annals of the Royal College of Surgeons of England* 2022;104:257-60.

13. Gerrard AD, Maeda Y, Miller J, Gunn F, Theodoratou E, Noble C*, et al.* Double faecal immunochemical testing in patients with symptoms suspicious of colorectal cancer. *British Journal of Surgery* 2023;110:471-80.
14. Godber IM, Todd LM, Fraser CG, MacDonald LR, Younes HB. Use of a faecal immunochemical test for haemoglobin can aid in the investigation of patients with lower abdominal symptoms. *Clinical Chemistry & Laboratory Medicine* 2016;54:595-602.
15. Johnstone MS, Burton P, Kourounis G, Winter J, Crighton E, Mansouri D*, et al.* Combining the quantitative faecal immunochemical test and full blood count reliably rules out colorectal cancer in a symptomatic patient referral pathway. *International Journal of Colorectal Disease* 2022a;37:457-66.
16. MacDonald S, MacDonald L, Godwin J, Macdonald A, Thornton M. The diagnostic accuracy of the faecal immunohistochemical test in identifying significant bowel disease in a symptomatic population. *Colorectal Disease* 2022;24:257-63.
17. Mowat C, Digby J, Strachan JA, McCann RK, Carey FA, Fraser CG*, et al.* Faecal haemoglobin concentration thresholds for reassurance and urgent investigation for colorectal cancer based on a faecal immunochemical test in symptomatic patients in primary care. *Annals of Clinical Biochemistry* 2021;58:211-9.
18. Mowat C, Digby J, Strachan JA, McCann R, Hall C, Heather D*, et al.* Impact of introducing a faecal immunochemical test (FIT) for haemoglobin into primary care on the outcome of patients with new bowel symptoms: a prospective cohort study. *BMJ Open Gastroenterology* 2019;6:e000293.
19. Nicholson BD, James T, East JE, Grimshaw D, Paddon M, Justice S*, et al.* Experience of adopting faecal immunochemical testing to meet the NICE colorectal cancer referral criteria for low-risk symptomatic primary care patients in Oxfordshire, UK. *Frontline Gastroenterology* 2019;10:347-55.
20. Nicholson BD, James T, Paddon M, Justice S, Oke JL, East JE*, et al.* Faecal immunochemical testing for adults with symptoms of colorectal cancer attending English primary care: a retrospective cohort study of 14 487 consecutive test requests. *Alimentary Pharmacology & Therapeutics* 2020;52:1031-41.
21. Tang A, Chandler S, Torkington J, Harris DA, Dhruva Rao PK. Adapting the investigation of patients on urgent suspected cancer pathway with lower gastrointestinal symptoms across Wales during COVID-19. *Annals of the Royal College of Surgeons of England* 2022;26:26.
22. Turvill J, Mellen S, Jeffery L, Bevan S, Keding A, Turnock D. Diagnostic accuracy of one or two faecal haemoglobin and calprotectin measurements in patients with suspected colorectal cancer. *Scandinavian Journal of Gastroenterology* 2018;53:1526-34.
23. Turvill J, Turnock D, Cottingham D, al. e. The Fast Track FIT study: diagnostic accuracy of faecal immunochemical test for haemoglobin in patients with suspected colorectal cancer. *Br J Gen Pract* 2021;71:E643–E51.
24. Withrow DR, Shine B, Oke J, Tamm A, James T, Morris E*, et al.* Combining faecal immunochemical testing with blood test results for colorectal cancer risk stratification: a consecutive cohort of 16,604 patients presenting to primary care. *BMC Medicine* 2022;20:116.
25. Archer T, Aziz I, Kurien M, Knott V, Ball A. Prioritisation of lower gastrointestinal endoscopy during the COVID-19 pandemic: outcomes of a novel triage pathway. *Frontline Gastroenterology* 2022;13:225-30.

26. Ayling RM, Lewis SJ, Cotter F. Potential roles of artificial intelligence learning and faecal immunochemical testing for prioritisation of colonoscopy in anaemia. *British Journal of Haematology* 2019;185:311-6.
27. Ball AJ, Aziz I, Parker S, Sargur RB, Aldis J, Kurien M. Fecal Immunochemical Testing in Patients With Low-Risk Symptoms of Colorectal Cancer: A Diagnostic Accuracy Study. *Journal of the National Comprehensive Cancer Network* 2022;20:989-96.e1.
28. Bujanda L, Sarasqueta C, Vega P, Salve M, Quintero E, Alvarez-Sánchez V*, et al.* Effect of aspirin on the diagnostic accuracy of the faecal immunochemical test for colorectal advanced neoplasia. *United European Gastroenterol J* 2018;6:123-30.
29. Cama R, Kapoor N, Sawyer P, Patel B, Landy J. Evaluation of 13,466 Fecal Immunochemical Tests in Patients Attending Primary Care for High- and Low-Risk Gastrointestinal Symptoms of Colorectal Cancer. *Digestive Diseases & Sciences* 2022;10:10.
30. Crooks C, Banerjea A, Jones J, Chapman C, Oliver S, West J*, et al.* Assessing empirical thresholds for investigation in people referred on a symptomatic colorectal cancer pathway: a cohort study utilising faecal immunochemical and blood tests in England. *medRxiv* 2023; 10.1101/2023.03.29.23287919:2023.03.29.23287919.
31. Georgiou Delisle T, D'Souza N, Tan J, Najdawi A, Chen M, Ward H*, et al.* Introduction of an integrated primary care faecal immunochemical test referral pathway for patients with suspected colorectal cancer symptoms. *Colorectal Disease* 2022a;08:08.
32. Juul JS, Hornung N, Andersen B, Laurberg S, Olesen F, Vedsted P. The value of using the faecal immunochemical test in general practice on patients presenting with non-alarm symptoms of colorectal cancer. *British Journal of Cancer* 2018;119(4):471-9.
33. Laszlo HE, Seward E, Ayling RM, Lake J, Malhi A, Stephens C*, et al.* Faecal immunochemical test for patients with 'high-risk' bowel symptoms: a large prospective cohort study and updated literature review. *British Journal of Cancer* 2022;126:736-43.
34. Maclean W, Limb C, Mackenzie P, Whyte MB, Benton SC, Rockall T*, et al.* Adoption of faecal immunochemical testing for 2-week-wait colorectal patients during the COVID-19 pandemic: an observational cohort study reporting a new service at a regional centre. *Colorectal Disease* 2021a;23(7):1622-9.
35. Morales Arraez D, Carrillo G, Adrian M, Gimeno Z, Quintero A. Role of faecal immunochemical testing in the diagnostic workup of patients with iron deficiency anaemia. *United Eur Gastroenterol J* 2018;6:A403–A4.
36. Mowat C, Digby J, Strachan JA, Wilson R, Carey FA, Fraser CG*, et al.* Faecal haemoglobin and faecal calprotectin as indicators of bowel disease in patients presenting to primary care with bowel symptoms. *Gut* 2016;65:1463-9.
37. Pin-Vieito N, Garcia Nimo L, Bujanda L, Roman Alonso B, Gutierrez-Stampa MA, Aguilar-Gama V*, et al.* Optimal diagnostic accuracy of quantitative faecal immunochemical test positivity thresholds for colorectal cancer detection in primary health care: A community-based cohort study. *United European Gastroenterology Journal* 2021;9:256-67.
38. Rodriguez-Alonso L, Rodriguez-Moranta F, Arajol C, Gilabert P, Serra K, Martin A*, et al.* Proton pump inhibitors reduce the accuracy of faecal immunochemical test for detecting advanced colorectal neoplasia in symptomatic patients. *PLoS One* 2018;13:e0203359.
39. Rodriguez-Alonso L, Rodriguez-Moranta F, Ruiz-Cerulla A, Arajol C, Serra K, Gilabert P*, et al.* The use of faecal immunochemical testing in the decision-making process for the endoscopic investigation of iron deficiency anaemia. *Clin Chem Lab Med* 2020;58:232-9.

40. Maclean W, Mackenzie P, Limb C, Zahoor Z, Whyte MB, Rockall T*, et al.* Diagnostic accuracy of point of care faecal immunochemical testing using a portable high-speed quantitative analyser for diagnosis in 2-week wait patients. *Colorectal Disease* 2021b;23:2376-86.
41. Tsapournas G, Hellström PM, Cao Y, Olsson LI. Diagnostic accuracy of a quantitative faecal immunochemical test vs. symptoms suspected for colorectal cancer in patients referred for colonoscopy. *Scandinavian Journal of Gastroenterology* 2020;55:184-92.
42. Sieg A, Thoms C, Lüthgens K, John MR, Schmidt-Gayk H. Detection of colorectal neoplasms by the highly sensitive hemoglobin-haptoglobin complex in feces. *International Journal of Colorectal Disease* 1999;14:267-71.
43. Hunt N, Rao C, Logan R, Chandrabalan V, Oakey J, Ainsworth C*, et al.* A cohort study of duplicate faecal immunochemical testing in patients at risk of colorectal cancer from North-West England. *BMJ Open* 2022;12:e059940.