# Efficacy and accuracy of faecal sampling by a digital rectal examination for FIT

<table>
<thead>
<tr>
<th>Journal</th>
<th>Annals of Clinical Biochemistry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>ACB-22-158.R2</td>
</tr>
<tr>
<td>Manuscript Type</td>
<td>Research Article</td>
</tr>
<tr>
<td>Date Submitted by the Author</td>
<td>02-Dec-2022</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Maclean, William; Royal Surrey NHS Foundation Trust, Colorectal Surgery; Benton, Sally; Surrey Pathology Services, Clinical Biochemistry; Surrey Pathology Services, Bowel Cancer Screening Southern Hub; Whyte, Martin; University of Surrey; Rockall, Timothy; Royal Surrey County Hospital NHS Foundation Trust, Minimal Access Therapy Training Unit (MATTU); Jourdan, Iain; Royal Surrey County Hospital NHS Foundation Trust</td>
</tr>
<tr>
<td>Keywords:</td>
<td>Clinical studies, Immunoassay, Laboratory methods, Haemoglobin, Analytes</td>
</tr>
</tbody>
</table>
Efficacy and accuracy of faecal sampling by a digital rectal examination for FIT

William Maclean¹, Sally C. Benton², Martin B. Whyte³, Timothy Rockall¹, Iain Jourdan¹

¹Colorectal Surgery, Royal Surrey NHS Foundation Trust, Guildford, UK
²Bowel Cancer Screening Hub at Royal Surrey NHS Foundation Trust, Guildford, UK
³University of Surrey, Guildford, UK

Corresponding author: Mr William Maclean, Research Fellow in General Surgery, Royal Surrey NHS Foundation Trust, Guildford, GU2 7XX, UK
Email: William.maclean@doctors.org.uk

DECLARATIONS

Competing interests: The authors declare no conflicts of interest with respect to the research, authorship, and/or publication of this article

Funding: Sentinel Diagnostics provided sampling devices and all reagents for the FOB Gold Wide®. The company was not involved in the design of study, the statistical analyses or the interpretation of data; GUTsfbc is a registered charity (1026791) that awarded a £500 grant towards the study. This money was subsequently not required and returned to the charity

Ethical Approval: Ethics for the project was approved by the London - South East Research Ethics Committee on 28th May 2019 (REC reference: 19/LO/0889, IRAS ID: 260384)

Guarantor: IJ

Contributorship: SCB, IJ and WM conceived the study. WM was the Principal investigator for the study and primary recruiter. WM wrote drafts of the manuscript. All authors reviewed and edited the manuscript and approved the final version

Acknowledgements: The authors wish to thank the research nurses at RSFT and in particular – Sinead Donlon, Celia Harris, Suzanna Tluk and Humyraa Aziz for their regular contribution in assistance for recruitment and sample delivery.
Abstract:

**Aim:** A digital rectal examination (DRE) during routine assessment for patients with abdominal symptoms provides an opportunity to obtain faeces from the glove for faecal immunochemical testing (FIT). Here, we compared sampling via DRE to the standard faecal sampling by patients.

**Method:** Patients were recruited to a prospective observational cohort study between July 2019 and March 2020. Patients provided a sample for the FOB Gold Wide® which was compared to a further sample taken at clinic via DRE. Clinicians reported whether they obtained a “good” sample filling all the grooves, a “poor” sample filling some of the grooves or no faecal sample. Cohen’s kappa was used to compare percentage agreement around a negative threshold of <10 μg haemoglobin/g of faeces. Sensitivity for serious bowel disease (SBD) was calculated.

**Results:** Of 596 patients who underwent attempted DRE sampling, there were 258 (43.3%) “good” samples, 117 (19.6%) “poor” samples and 221 (37.1%) with no sample to wipe in the grooves. Cohen’s kappa dropped from 0.70 to 0.30 for the “good” and “poor” samples respectively. Of those with DRE samples and definitive diagnostic outcomes, the sensitivity for SBD dropped significantly from 76.0% to 41.7% between “good” and “poor” samples respectively (p=0.041).

**Conclusions:** A “good” sample obtained by DRE provides comparable results to samples obtained by patients. This creates potential benefit in speed and ease of testing for patients. However, not all DRE sampling attempts are successful, and the clinician must be satisfied that enough faeces is obtained to wipe adequately into all grooves.
Introduction:

Speed of obtaining a faecal haemoglobin (f-Hb) result could be improved by use of point-of-care (POC) devices or the use of high-speed laboratory-based devices with online reporting, but this does not solve the initial difficulty of obtaining the faecal sample. Some patients may find collecting a sample for faecal immunochemical testing (FIT) to be challenging. Previously identified reasons for non-compliance in faecal sampling include embarrassment, fear of results, concerns around hygiene and contamination, discretion and privacy, and lack of information.¹ Return rates for symptomatic patients reported by Chapman et al for a primary care pathway were reported as 91.4% and 1.3% of samples were unable to be analysed.² Furthermore, there is potential for delay as the patient needs to obtain the FIT sample collection device, collect their faeces, and deliver it back to the requesting clinician or laboratory. A faecal sample obtained by digital rectal examination (DRE) has the potential to address these issues.

Two studies have reported the diagnostic accuracy of obtaining faeces via DRE and wiping the faeces from the glove onto the FIT sampling stick.³⁴ A DRE is recognised as a routine part of the examination of patients with bowel symptoms, so the process of obtaining the sample does not subject the patient to any extra steps. The responsibility is moved from the patient to clinician and can be taken at first consultation. However, obtaining faeces via DRE has not been recommended by the manufacturers of the FIT analysers. Previous studies have compared the diagnostic accuracy of DRE sampling to patient collected samples with guaiac faecal occult blood testing in screening patients. These have shown differing levels of accuracy and completion rates.⁵⁻⁷ There are no comparative studies of paired samples for FIT.

The aims of this study were to determine the feasibility of using DRE as a sampling method for FIT and to compare the f-Hb results and diagnostic accuracy with the standardised method of sampling. The objective was to obtain paired samples of a FIT faecal sample via DRE in clinic with comparison to a home collected sample that the patient brought to the clinic.
Methods:

Design and intervention

A prospective observational cohort study was designed in line with the updated STARD checklist for reporting diagnostic accuracy studies. Recruitment took place between July 2019 and March 2020 from symptomatic patients referred to the Royal Surrey NHS Foundation Trust (RSFT). These patients were invited to the “POC FIT” study (REC: 19/LO/0889) by post and asked to provide two samples from the same bowel motion/faecal sample. One sample was analysed with the POC QuikRead go® (QRG), the other with the laboratory-based FOB Gold Wide® (Sentinel Diagnostics, Italy). The diagnostic accuracy of the QRG and its comparison to laboratory-based testing has previously been reported for this study. A third sample was obtained via DRE and also analysed with the laboratory-based FOB Gold Wide® to allow direct comparison with the sample collected by the patient.

The postage invitation supplied to the patients contained the specialised collection devices suitable for both the QRG and FOB Gold Wide®. The patient information leaflet explained the technique for home faecal sampling and the plan to obtain a third sample via DRE in clinic. They were asked to bring their samples to their clinic appointment, where written consent was obtained if they agreed to enrol in the observational study.

Inclusion and Exclusion

All patients ≥18 years of age referred on the two-week wait (TWW) pathway were eligible. This pathway consisted of patients referred from primary care to colorectal surgery at RSFT with ‘red flag’ bowel symptoms or anaemia according to NG12 guidance. Patients had to have capacity to consent and have provided their samples from fresh faeces and not a stoma bag. Patients who had collected their FIT samples more than 10 days prior to clinic appointment were excluded - based upon sample stability recommendations by the manufacturer.

Specimen collection and analysis

Three colorectal doctors were involved in the recruitment from their clinics and collected faecal samples via DRE from those that consented to the study. Faeces from the glove was
wiped directly onto the grooves of the FIT sample stick of the collection device for the FOB Gold Wide®. Clinicians made a judgement on whether any faecal matter was obtained via DRE or not. If none was obtained, a sample was not sent. If a sample was obtained, this was categorised into: “good” sample, filling all the grooves of the sampling stick; or “poor” sample, filling some of the grooves. Patients and clinicians were blinded to f-Hb results in order not to influence decision-making for standard diagnostic tests. All FOB Gold Wide® samples were analysed in the research laboratory based at the Bowel Cancer Screening Southern Hub (BCSH) in Guildford. All laboratory analysis was overseen by a state-registered biomedical scientist. The FOB Gold Wide® at the BCSH is currently used for research purposes and not routine clinical use and therefore is not accredited, however it has undergone thorough validation. All assays that are routinely used at RSFT bowel cancer screening hub and biochemistry department are UKAS accredited. Further details on FIT methods for this study can be found in the FITTER checklist as supplied in a supplementary file.

Result interpretation and statistical analysis

The limit of detection and quantification for the FOB Gold Wide® are <3 μg of haemoglobin per gram of faeces (μg/g) and >1700 μg/g. For comparison of paired samples, results outside these ranges were excluded and subsequent non-paired data points were also excluded. The relationship of these pairs was demonstrated by a Bland-Altman plot. Results where f-Hb <10 μg/g was considered a negative result. The proportion of results that showed f-Hb <10 μg/g was compared between the two systems using Chi squared testing. Percentage agreement at this threshold was calculated and assessed with Cohen’s kappa coefficient.12

Diagnostic accuracy was determined by comparing f-Hb results with eventual definitive diagnostic outcomes through colonoscopy or CT colonography. Patients undergoing flexible sigmoidoscopy were only included if they had presented with perianal symptoms or anorectal bleeding (bright red blood seen separate to the faeces in the pan or on the paper). Sensitivity and specificity for CRC and serious bowel disease (SBD) was determined using the threshold of 10 μg/g. SBD was defined as either a diagnosis of CRC, inflammatory bowel disease or high-risk adenoma (HRA). Patients were defined as having HRAs if they were...
found to have an advanced adenoma (≥10mm, any sessile serrated lesion or adenomas that contained high risk dysplasia), or to have ≥5 polyps, as per the British Society of Gastroenterology guidelines for surveillance. All CRC diagnoses were confirmed by histology reports.

Data
A secure web-based clinical database, associated with RSFT and approved by the local information governance team, was developed to audit colorectal patients on the urgent referral pathway. Recruited participants for the study were assigned a unique trial number and their FIT results were entered into the database in a pseudonymised fashion. Access to this data was restricted to members of the research team.

Results:
There were 629 patients consented and included for the POC FIT study. There were 596 patients where DRE sampling was attempted. Figure 1 shows the patient pathway for inclusion and categorisation for those that went on for DRE sampling.

[Insert Figure 1.]

Of 596 patients who underwent DRE (table 1), there were 221 (37.1%) DREs where no faecal matter was obtained on the glove for sampling. When sampling was achieved (in 375 patients), the clinician graded the sample as “good” in 258 of the cases and “poor” in 117 of the cases.
Table 1: Age and sex of 596 patients where DRE sampling was attempted

<table>
<thead>
<tr>
<th>Sex</th>
<th>Male</th>
<th>Female</th>
<th>18-39</th>
<th>40-49</th>
<th>50-59</th>
<th>60-69</th>
<th>70-79</th>
<th>80-89</th>
<th>≥90</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>All DRE attempts as of total</td>
<td>290 (48.7%)</td>
<td>306 (51.3%)</td>
<td>18 (3.0%)</td>
<td>37 (6.2%)</td>
<td>117 (19.6%)</td>
<td>154 (25.8%)</td>
<td>174 (29.2%)</td>
<td>90 (15.1%)</td>
<td>6 (1.0%)</td>
<td>596</td>
</tr>
<tr>
<td>“Good” samples (% within category)</td>
<td>119 (41.0%)</td>
<td>139 (45.4%)</td>
<td>4 (22.2%)</td>
<td>17 (45.9%)</td>
<td>36 (30.8%)</td>
<td>61 (39.6%)</td>
<td>89 (51.1%)</td>
<td>50 (55.6%)</td>
<td>1 (16.7%)</td>
<td>258</td>
</tr>
<tr>
<td>“Poor” samples (% within category)</td>
<td>61 (22.4%)</td>
<td>52 (17.0%)</td>
<td>4 (22.2%)</td>
<td>5 (13.5%)</td>
<td>23 (19.7%)</td>
<td>28 (18.2%)</td>
<td>42 (24.1%)</td>
<td>13 (14.4%)</td>
<td>2 (33.3%)</td>
<td>117</td>
</tr>
<tr>
<td>No sample (% within category)</td>
<td>106 (36.6%)</td>
<td>116 (39.7%)</td>
<td>10 (55.6%)</td>
<td>15 (40.5%)</td>
<td>58 (49.6%)</td>
<td>65 (42.2%)</td>
<td>43 (24.7%)</td>
<td>27 (30.0%)</td>
<td>3 (50.0%)</td>
<td>221</td>
</tr>
</tbody>
</table>

Of all the DRE samples obtained, the proportion that gave a f-Hb result of <10 μg/g was 81.3% (305/375), this compares to 80.3% (301/375) from the standard sampling method by the patient and the difference was not significant (p=0.71). There were 211/258 (81.7%) “good” samples with a f-Hb <10 μg/g vs 207/258 (80.2%) from their paired standard samples which was not statistically different (p=0.65). There were 94/117 (80.3%) “poor” samples with a f-Hb <10 μg/g, which was the same proportion as seen from their paired standard samples.

The agreement at a threshold of <10 μg/g for all the paired samples was 86.6% and Cohen’s kappa coefficient was 0.57, demonstrating moderate agreement. However, the agreement was 90.7% for the “good” samples with a Cohen’s kappa coefficient of 0.70 – demonstrating substantial agreement. The agreement was 77.7% for the “poor” samples with a Cohen’s kappa coefficient of 0.30 for the “poor” samples demonstrating fair agreement.

After exclusion of pairs with a result below the limit of detection and above the limit of quantification, the f-Hb comparisons are demonstrated in figure 2 as a Bland-Altman plot. In this graph, there was a total of 69 paired f-Hb samples, these are plotted using the mean of the f-Hb values vs the percentage difference. The bias was -43.1 indicating that the DRE
sample tended to give a lower f-Hb concentration. The dotted lines represent the upper and lower limits of agreement (153.0 and -239.3 respectively), and the filled line is the bias. The graph uses a log-scale for the x-axis.

[Insert Figure 2.]

Diagnostic accuracy

Of the 629 patients included in the study, 553 underwent colonic investigations that were suitable for comparing f-Hb results with diagnostic outcomes. For the 375 patients where a faecal sample was obtained at DRE, 342 obtained suitable colonic investigations. There were 14 CRCs diagnosed in the study (Table 2). All of these had an f-Hb >10 \( \mu g/g \) when samples were taken by the patient. Ten of these 14 patients had a DRE where a faecal sample was obtained and nine of these were described as a “good” sample covering all the grooves. All ten DRE obtained samples had an f-Hb >10 \( \mu g/g \). Comparison in diagnostic accuracy between the two sampling methods for CRC is displayed in figure 3.

[Insert Figure 3 and Table 2.]

Table 2: f-Hb results of the patients diagnosed with colorectal cancer

<table>
<thead>
<tr>
<th>f-Hb from DRE sample method (( \mu g/g ))</th>
<th>f-Hb from normal sample method (( \mu g/g ))</th>
<th>Difference (( \mu g/g ))</th>
<th>DRE sample description</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td>356</td>
<td>-295</td>
<td>good</td>
</tr>
<tr>
<td>81</td>
<td>302</td>
<td>-221</td>
<td>good</td>
</tr>
<tr>
<td>167</td>
<td>302</td>
<td>-135</td>
<td>good</td>
</tr>
<tr>
<td>601</td>
<td>1040</td>
<td>-439</td>
<td>good</td>
</tr>
<tr>
<td>764</td>
<td>1515</td>
<td>-751</td>
<td>poor</td>
</tr>
<tr>
<td>871</td>
<td>&gt;1700</td>
<td>negative</td>
<td>good</td>
</tr>
<tr>
<td>977</td>
<td>116</td>
<td>861</td>
<td>good</td>
</tr>
<tr>
<td>1638</td>
<td>169</td>
<td>1469</td>
<td>good</td>
</tr>
<tr>
<td>&gt;1700</td>
<td>592</td>
<td>positive</td>
<td>good</td>
</tr>
<tr>
<td>&gt;1700</td>
<td>471</td>
<td>positive</td>
<td>good</td>
</tr>
<tr>
<td>n/a</td>
<td>43</td>
<td>n/a</td>
<td>no faeces</td>
</tr>
</tbody>
</table>
There were 52 diagnoses of other SBDs (CRC, inflammatory bowel disease or HRA). For samples taken by the patient, the sensitivity and specificity for SBD was 65.4% and 87.6% respectively. 37/52 patients had a DRE where a faecal sample was obtained. Comparison in diagnostic accuracy between the two sampling methods for SBD is displayed in figure 4. Of the 37 patients with a DRE faecal sample, 25 were described as “good” and 12 were described as “poor”. Of the “good” samples, 6/25 were <10 μg/g (sensitivity 76.0%) and of the “poor” samples, 7/12 were <10 μg/g (sensitivity – 41.7%). Using Chi squared to test differences in proportions, this difference in sensitivity between “good” and “poor” samples was significant (p=0.041). The specificities for the “good” and “poor” samples were 88.5% and 82.3% respectively (p=0.139).

Discussion:
This is the first patient comparative study of paired f-Hb results through the different sampling techniques for FIT. When the clinician was satisfied there was a good volume of faeces to wipe in the grooves of the FOB Gold Wide® sample collection device, the results were reliable showing substantial agreement between the sampling methods (Cohen’s kappa 0.70) and high sensitivity for SBD (76.0% from DRE sampling vs 65.4% for standard patient sampling technique). However, if the clinician was not satisfied with the sample, the agreement was mild (Cohen’s kappa 0.30) and the sensitivity for SBD significantly dropped to only 41.7%. In addition, the overall success of obtaining any form of sample was achieved in only 62.9% of cases. In patients with CRC and for whom the clinician had obtained faeces for sampling by DRE, none out of the ten had a false negative result using a threshold of 10 μg/g. However, in these cases, a “good” sample was obtained in all but one of the patients.

For patients referred on the TWW in the United Kingdom there is a requirement for patients to be seen by a specialist within two weeks from referral and for a diagnosis to be
established within 28 days. If DRE were to be combined with POC testing, which has been shown to be comparable in sensitivity to laboratory-based testing\(^9\), then there is potential for a clinician to obtain a sample and have a result all within an initial consultation. This could better aid decision-making regarding referral or investigation choice. It would also help patient avoid the need to obtain the sample at home and the potential anxiety of waiting for results.

The quick turnaround of the DRE obtained sample can channel a more rapid decision for colonic imaging by endoscopy or CT colonography by avoiding delay in waiting for return of a sample by the patient. The high specificities seen within our cohort via both forms of sampling shows that FIT is a reliable means to triage patients. Obtaining f-Hb results quicker and with better compliance through DRE may increase the use of FIT prior to referral and thus impact decision-making to better prioritise those with f-Hb ≥ 10 µg/g. This will help fast-track more at-risk patients, whilst minimising those undergoing unnecessary invasive colonic investigations creating a more cost-effective pathway. However, if the application of FIT was applied to a population with low prevalence of the condition, the positive predictive value of the test would necessarily fall, and greater numbers of patients would be referred on to the TWW pathway with false positive results. This may result in increased patient anxiety and increased resource pressures.

This data on DRE sampling for FIT has shown that faeces collected in this way is comparable to patient own sampling in a symptomatic cohort provided that the clinician feels that they have obtained an adequate sample. As there was a significant drop in sensitivity for SBD when a sample was deemed “poor”, we would only recommend processing “good” samples for clinical application. A survey from Boston showed that DRE sampling would be acceptable to the majority of both patients and clinicians, but this was dependent on completion rates.\(^{14}\) Initially, they found that 54% of care providers would routinely offer DRE sampling for FIT, but this would go up to 88% if they found the results to be comparable to patient collected samples and if the completion rate was 75%.\(^{14}\) Our study found only 258 of the 596 DRE sampling attempts obtained were “good” and therefore this practice cannot be relied upon for most cases. However, if DRE and POC FIT were to be used at first consultation and a FIT sample collection device given to all patients where a good sample
was not obtained, then this would accelerate the pathway for over 40% of patients where immediate decisions could be made at first primary care consultation.

FIT has rapidly taken off for symptomatic patients and the application is not solely for low-risk patients as a rule in investigation as per the NICE guidance from 2017\textsuperscript{15}. The results of the NICE FIT study demonstrated the high sensitivity of FIT for high-risk patients in primary care as a rule out investigation\textsuperscript{16} and FIT has been shown to be safe and effective for TWW patients in the coronavirus pandemic to better allocate diminished endoscopy capacity\textsuperscript{17,18}. Therefore, we suggest that FIT via DRE could also be applied in primary care as a rule out investigation for potential TWW patients. Furthermore, there is interest growing for repeat FIT to further improve sensitivity\textsuperscript{19–21} - in this situation, DRE sampling may also be the method of choice to obtain the additional test. If a routine referral to secondary care has been made due to a negative FIT from primary care, the specialist may find a second sample via DRE helpful to further determine the need for invasive colonic investigation.

**Study Limitations**

There are different sample collection devices for the different FIT systems and therefore a larger study is required to assess the comparison of results and diagnostic accuracy of all these. The design of the sampling stick for the FOB Gold Wide\textsuperscript{®} and the QRG are similar with circular grooves at the tip to collect the faeces. One of the DRE sampling diagnostic accuracy studies used the HM-JACKarc system.\textsuperscript{3} The sampling stick for this has a single small groove and therefore may be more appropriate to wipe the smaller amounts of faeces than the multiple circumferential grooves of the FOB Gold Wide\textsuperscript{®} and QRG.

The evidence presented here is based on DREs performed by only three clinicians and the patient cohort was from a single centre. The ability to obtain faeces from a DRE may be different according to experience from other colorectal clinicians or general practitioners in primary care.

**Conclusions:**
DRE sampling produces FIT results comparable to patient own sampling when the clinician can obtain a “good” sample. This creates potential benefits in terms of speed and ease of testing which could accelerate the patient pathway. However, not all DRE attempts are successful, and the clinician must be sensitive to the possibility of poor sampling and ensure that in such cases patients are offered standard home testing with FIT.

References:


safety-netting of patients with symptoms and low faecal haemoglobin concentration

20. Farkas NG, Fraser CG, Maclean W, et al. Replicate and repeat faecal immunochemical
tests in symptomatic patients: A systematic review. Ann Clin Biochem. 2022
May;45632221096036.

testing in patients at risk of colorectal cancer from North-West England. BMJ Open.
Figure 1: Patient Pathway and comparative results of DRE with normal patient sampling technique
Figure 2: Bland-Altman plot to show the relationship between the 69 paired samples for both “good” and “poor” DRE samples vs patient own sampling.
Figure 3: 2X2 tables to compare the diagnostic accuracies for colorectal cancer with patient own collected samples vs DRE collected samples

**Patient own samples (n=553)**

<table>
<thead>
<tr>
<th></th>
<th>Colorectal Cancer</th>
<th>No Colorectal Cancer</th>
<th>Positive predictive value</th>
<th>Negative predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>f-Hb ≥10 µg/g</td>
<td>14</td>
<td>82</td>
<td>17.1%</td>
<td>93%</td>
</tr>
<tr>
<td>f-Hb &lt;10 µg/g</td>
<td>0</td>
<td>457</td>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100%</td>
<td>Specificity</td>
<td>84.8%</td>
<td></td>
</tr>
</tbody>
</table>

**DRE collected samples (n=342)**

<table>
<thead>
<tr>
<th></th>
<th>Colorectal Cancer</th>
<th>No Colorectal Cancer</th>
<th>Positive predictive value</th>
<th>Negative predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>f-Hb ≥10 µg/g</td>
<td>10</td>
<td>55</td>
<td>15.4%</td>
<td>90%</td>
</tr>
<tr>
<td>f-Hb &lt;10 µg/g</td>
<td>0</td>
<td>277</td>
<td></td>
<td>100%</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>100%</td>
<td>Specificity</td>
<td>83.4%</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4: 2X2 tables to compare the diagnostic accuracies for serious bowel disease with patient own collected samples vs DRE collected samples

<table>
<thead>
<tr>
<th></th>
<th>Patient own samples (n=553)</th>
<th>DRE collected samples (n=342)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Serious bowel disease</td>
<td>No serious bowel disease</td>
</tr>
<tr>
<td>f-Hb ≥10 μg/g</td>
<td>34</td>
<td>62</td>
</tr>
<tr>
<td>f-Hb &lt;10 μg/g</td>
<td>18</td>
<td>439</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>65.4%</td>
<td>64.9%</td>
</tr>
<tr>
<td>Specificity</td>
<td>87.6%</td>
<td>86.6%</td>
</tr>
</tbody>
</table>

Positive predictive value = 35.4%
Negative predictive value = 96.1%
Positive predictive value = 36.9%
Negative predictive value = 95.3%
Appendix - FITTER Checklist

Specimen collection and handling

Patients invited to the POC FIT study were sent packs (832 patients). Postage was sent using Royal Mail first-class. The pack contained the appointment details and the FIT collection devices for FOB Gold Wide® (Sentinel Diagnostics, Italy) and the QuikRead go® (Aidian Oy, Espoo, Finland). The pack also contained the appointment letter, an instruction leaflet for the sample to be collected by the patient and a patient information leaflet.

The two tubes for each patient were pre-labelled with their allocated pseudonymised number. They were placed in a sealable sample bag which was used for hygienic transport for patients to bring the sample to clinic with them. The sample bag had a space requesting date the samples were taken for the patient to write on. The specimens from the stool were requested in the immediate time from passing the stool to avoid haemoglobin degradation.

Patients were met in their two-week wait clinic and if they had provided the two samples, were offered participation in the trial and written consent was obtained. The third digital rectal examination samples were obtained through examining the patient in the left lateral position. The finger sweep in the rectum attempted to transfer stool to the glove of the examiner, any stool removed was wiped into the grooves of the FOB Gold Wide® sampling stick. If any stool was available to be wiped into the grooves, the clinician documented the category of the sample to be “good” or “poor”. The sample was labelled with the pseudonymised patient details, marked as a DRE sample and the date collected.

FOB Gold Wide® Analyser (Sentinel Diagnostics, Italy)

The FIT samples for the FOB Gold Wide® Analyser were collected from patients at their clinic appointment. The DRE samples were placed in the same bag for each patient. Samples were then kept in a sample fridge before being hand delivered to the BCSP Southern Hub, Surrey Research Park, Guildford for analysis.
Each FOB Gold Wide® sample vial contained 1.7 ml of stabilising buffer. The test stick with grooves delivered the set volume of stool to be suspended in the buffer solution. Excess stool was prevented going into the buffer by the funnel system within the tube.

**Analysis**

Analysis was carried out under supervision of a state registered biomedical scientist on the automated SENTiFIT 270 analyser using the FOB Gold Wide® method (Sentinel Diagnostics, Italy) according to manufacturer protocols. Results were determined as nanograms of haemoglobin per millilitre of buffer within the sample tube (ng/ml). From this the haemoglobin concentration was calculated and reported as micrograms of haemoglobin per gram of faeces (µg/g). The analyser works on the conversion where 1 ng/ml is equivalent to 0.17 µg/g. Each sample was analysed once.

- Analytical working range: 3-170 µg Hb/g faeces
- Whether samples outside the range were diluted (factor) and re-assayed: diluted by the analyser 1/10 to extend the analytical range to 1,700 µg Hb/g faeces
- Source of calibrator(s): supplied by Sentinel Diagnostics
- Number of calibrator(s): 1 calibrator diluted by the analyser to 6 concentrations
- How calibrator concentrations were assigned: N/A
- Calibration process including frequency: Calibration is carried out when the Lot number of the latex changes, every 28 days, when the QC values are not within the acceptable range, and after certain maintenance support procedures.
- Analytical imprecision, ideally with number of samples analysed, concentrations, and mean, SD and CV: the imprecision quoted by the manufacturer is: within-run (20 replicates for 3 runs at 8 concentrations) CV ≤7% for samples ≤ 90 ng/ml or SD ≤ ± 5 ng/ml, samples > 90 ng/ml CV ≤ 7.0%; between run (21 days, 2 runs/day2 tests/day for 6 concentrations CV ≤ 7% or SD ≤ ± 5 ng/ml, samples > 90 ng/ml CV ≤ 7.0%.
- The number, training and expertise of the persons performing the analyses and recording the results: 3 staff members trained; 1 state registered Healthcare Scientist, 2 research assistants, all with experience of running analysers for FIT.

**Quality Management:**
• Source (address) or description of internal quality control materials: Sentinel Diagnostics, SENTINEL CH. SpA, 20152 Milano - Italy
• Number of controls: 2
• Assigned target concentrations and ranges: 3 lot numbers of QC used during the study, QC level 1 means approximately 85 ng/ml, range ± 4 ng/ml, QC level 2 means approximately 300 ng/ml, range ± 7 ng/ml.
• How target concentrations were assigned: 20 data points over multiple days/runs. Mean and ± 2 SD’s calculated for acceptable range
• Rules used for acceptance and rejection of analytical runs. Westgard Rules: 2 QCs within ± 2SD’s at the beginning and the end of the run.
• Participation in external quality assessment schemes: (name and address of scheme): NEQAS for Faecal Haemoglobin, UK; WEQAS for Faecal Haemoglobin, UK
• Frequency of EQA challenges: Once a month
• EQA performance attained: Acceptable performance
• Accreditation held by the analytical facility (address): Routine service assays carried out in the bowel cancer screening hub and the biochemistry department at Royal Surrey NHS Foundation Trust are all UKAS accredited under ISO 1587. As the two assays in this study are not used for routine service - neither have accreditation. However, FOB Gold Wide® has undergone thorough validation.

Data handling
The f-Hb results were recorded electronically and uploaded prospectively into the secure web-based clinical database for the study. Laboratory staff were not involved in clinical decision-making and were therefore blinded to patient presentations and outcomes. Time between sample taken and sample processing was automatically calculated and displayed to ensure samples were processed within 10 days of collection.

Result Interpretation
The FIT results were compared for correlation and with definitive diagnostic outcomes as described in the Methods section of the manuscript. Cut-off for a negative result was set at f-Hb <10 μg/g.