PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Priority stratification for colonoscopy based on 2-sample faecal immunochemical test screening: results from a cross-sectional study at an endoscopy clinic in Japan
AUTHORS	Toyoshima, Osamu; Yamaji, Yutaka; Nishizawa, Toshihiro; Yoshida, Shuntaro; Yamada, Tomoharu; Kurokawa, Ken; Obata, Miho; Kondo, Ryo; Toba, Masahito; Koike, Kazuhiko

VERSION 1 – REVIEW

REVIEWER	Smith, Steve University Hospitals of Coventry & Warwickshire NHS Trust, Midlands & NW Bowel Cancer Screening Hub, Hospital of St. Cross
REVIEW RETURNED	03-Jan-2021

GENERAL COMMENTS	This is a highly topical subject and this paper has the potential to contribute significantly to the current discussions regarding the use of FIT. As the authors point out the use of FIt may have considerable benefit in prioritising colonoscopy resources especially during the COVID-19 pandemic.. Unfortunately this study as currently presented has too many unanswered questions to be of value. The authors do not provide any information regarding the FIT method(s) used in the study. It is not possible therefore to evaluate what is meant by a positive FIT result. Were the methods used purely qualitative or were they quantitaive or mixture? What were the thresholds above which a person was deemed to have a positive FIT? The authors recognise this in the limitations of the study but they don't even quote anything about the method used in their own institution. A recent paper suggests knowing the methods used is important see Chapman et al 2020 (below) There is no indication of the source of the FIT positive patients. How many were from people screened with a FIT test, how many patients had symptoms before doing the FIT test or were surveillance patients. In order to understand the results better this is important information which is lacking. For the FIT (+) positive patients the authors do not indicate how many people did two tests with one positive and one negative and how many people just did the one positive test. It is possible that this latter group should have been in the FIT (2+) group if they had done the second test. In the discussion the authors suggest that FIT is better at detecting the more advanced cancers rather than the earlier lesions. I think that there is eveidence to contradict this which they have failed to discuss (Moss et al 2017 and Clark et al 2020 below) Chapman C, Banerjea A. et al clin chem Lab Med 2020 Oct 29;/j/cclm.ahead-of-print/cclm-2020-1170.xml. doi10.1515/cclm-2020-1170 . Choice of faecal immunochemical test matters:

	comparison of OC-Sensor and HM-Jackarc, in the assessment of patients at high risk of colorectal cancer. Moss S., Matthews C., et al GUT 2017:66:1631-44 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. Clark G., Strachen JA., et al GUT 2020 doi 10.1136/gutjnl-2019-320297 . Transition to quantitative faecal immunochemical resting from guaiac faecal occult blood testing in a fully rolled-out population based national bowel screening programme.

REVIEWER	Bardou, Marc Centre Hospitalier Universitaire de Dijon, CIC-P INSERM 1432
REVIEW RETURNED	18-Jan-2021

GENERAL COMMENTS	Revue on MS bmjopen-2020-046055 by Osomu Tosyoshimada and colleagues. In this paper authors assessed priority stratification for colonoscopy based on the fact that patients had one or two positive FIT. They found that Detection rates of all (and invasive) cancers in the FIT (2+), FIT (+), and non-FIT groups were 12.1% (8.3%), 1.9% (0.5%), and 0.4% (0.2%), respectively. Although the paper may be of interest I have some significant concerns with it in its present form. The first concern I have is that the clinical purpose of the paper is not obvious to me. I wonder if authors want to suggest that CRC screening through the combination of 2 FIT has to become the standard approach in all countries with a FIT-based organized CRC screening program (CRCSP), or that in the case of limited access to endoscopy facilities, those with two FIT+ should be prioritized versus those with only one. Many countries have CRCSP that are based on a single FIT result, and I don't think the paper is likely to make them change. Indeed the purpose of a CRCSP is not to detect advanced neoplasia, but preneoplastic or early stage lesions. Here authors suggest that 2 FIT+ is more specific than 1 FIT+, whereas a CRCSP has to balance sensitivity and specificity. To be more specific. I first have a concern with the FIT conducted in patients bellow the age of 50, but it may be because of lack of knowledge of the way the Japanese CRCSP is organised. Indeed in many countries CRC screening only starts at 50. It seems to me, from my reading of table 1, that some patients included on this study had a FIT even bellow the age of 40. Can authors comment on that? I think authors should rerun their analysis only including those with 2 FIT results and comparing FIT+/+ vs FIT+/- (FIT+/- being for +/- or -/+). Mixing FIT+ (i.e. those who had only one test which was

	positive) and FIT+/- to make a single group of one single FIT is way too confusing, particularly when haemoglobin level is not provided. It would also have been useful to run the analysis based on Hb threshold, because it is well known that the lower it is, the highest is the likelihood of diagnosing non-neoplastic lesion such as adenomas. In this extent, even if authors are unable to do this because they have no idea of HB thresholds that have been used, which is a major pitfall, it would be useful to be provided with data on non-neoplastic findings. As it is to have the proportion of patients who were FIT+/+ and FIT+/-, and FIT-/- among all those who had two FIT results, stratified on age. I don't think the comparison with the group of patients who underwent colonoscopy is valid, at least the way it has been done, as it is a mix of different level of risk, mostly for the surveillance group which can be really heterogeneous. Other concerns The "Ethics paragraph sound odd" Authors write: "This study was approved by the Certificated Review Board, Hattori Clinic on September 6th, 2019 (approval no. S1909-U06, registration no. UMIN000018541). Written informed consent was obtained from the patients. All clinical investigations were conducted according to the ethical guidelines of the Declaration of Helsinki". But to my understanding this study consists on the retrospective analysis of data collected as part of an organised CRCSP. Same for the patient and public involvement paragraph which states that "Patients and/or the public were involved in the design, or conduct, or report, or dissemination plans of this research." How is that so? The discussion goes to fast on the limitations of the present study.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Dr. Steve Smith, University Hospitals of Coventry & Warwickshire NHS Trust, Midlands & NW Bowel Cancer Screening Hub, Hospital of St. Cross Comments to the Author:
This is a highly topical subject and this paper has the potential to contribute significantly to the current discussions regarding the use of FIT. As the authors point out the use of FIt may have considerable benefit in prioritising colonoscopy resources especially during the COVID-19 pandemic.. Unfortunately this study as currently presented has too many unanswered questions to be of value.
The authors do not provide any information regarding the FIT method(s) used in the study. It is not possible therefore to evaluate what is meant by a positive FIT result. Were the methods used purely qualitative or were they quantitaive or mixture? What were the thresholds above which a person was deemed to have a positive FIT? The authors recognise this in the limitations of the study but they don't even quote anything about the method used in their own institution. A recent paper suggests knowing the methods used is important see Chapman et al 2020 (below) There is no indication of the

source of the FIT positive patients. How many were from people screened with a FIT test, how many patients had symptoms before doing the FIT test or were surveillance patients. In order to understand the results better this is important information which is lacking.
[Ref A] *Chapman C, Banerjea A. et al clin chem Lab Med 2020 Oct 29;/j/cclm.ahead-of-print/cclm-2020-1170.xml. doi10.1515/cclm-2020-1170 . Choice of faecal immunochemical test matters: comparison of OC-Sensor and HM-Jackarc, in the assessment of patients at high risk of colorectal cancer.*

Thank you for your insightful comment. We appreciate your helpful suggestion for this article.

Because the guidelines for colorectal cancer screening in Japan only recommend the 2-sample FIT as standard with no specific kits or cutoff values, we suspected the kits used and cutoff values to be various in our patients, which we were not informed of in many cases.

Amongst the 1282 FIT-positive patients in this study, 14 (1%) underwent FIT screening at our institute. We used OC-Auto Sampling Bottle 3 (Eiken Chemical Co., Ltd., Tokyo, Japan) with the threshold of 32 µg haemoglobin/g faeces.

As differences in FIT kit features and thresholds have been known to affect screening performance, these variations are certainly a limitation of our study. However, the discordance in FIT results between different days is remarkable; therefore, the 2-sample FIT was adopted. A notable difference in the results between 2-positive and 1-positive FIT groups shown in our study suggests a common trend kits brand.

All patients with positive FIT results were primarily screened with FIT. No patient underwent FIT to make decisions in response to their symptoms or history of polypectomy. Among FIT-positive patients, 31 had symptoms and 19 had a history of colorectal lesions. Noninvasive carcinomas were found in three patients with 2-positive FITs and haematochezia. No cancer was detected in patients with positive FIT and history of colorectal lesions.

The following sentence was added to **Study design** of Methods section. (Page7, Line10 in the original manuscript)

The FIT kits included both qualitative and quantitative types. The FIT kit brand and cutoff values for positivity were chosen by the institutes conducting the FIT. At our institute, FIT was performed using OC-Auto Sampling Bottle 3 (Eiken Chemical Co., Ltd., Tokyo, Japan) with the threshold of 32 µg haemoglobin/g faeces.

Then, the following sentence (Page7, Line13-14 in the original manuscript) was changed and moved before the sentences above.
From:
The results of FIT were based on the test conducted at our clinic or at the referral medical institutions.
To:
FITs were conducted at our clinic or at referral medical institutions.

We changed the second paragraph of the **Limitations and Strengths** and additionally cited the reference above.
From:
Second, since our institute is specialised in endoscopies, many patients were referred from other medical institutions for colonoscopy. The category of FIT (+) included various categories of positivity for FIT: 1-positive result in 2 samples, 1-positive result in 1 sample, unknown number-positive results in 2 samples, and so on. The brand names of FIT kits or cutoff values for positivity were also unknown in many cases. However, a similar trend was seen when patients with 1-positive result in 2 samples from the FIT (+) group were separately analysed (**Supplementary Table 1**).
To:
Second, the FIT kit brands and cutoff values for positivity were various and unknown in many cases that were referred from other medical institutions for colonoscopy. The guidelines for colorectal cancer screening in Japan only recommend the 2-sample FIT as standard, with no specific kits or cutoff values. As differences in FIT kit features and thresholds have been known to affect screening performance,[Ref A] these variations are certainly a limitation of our study. However, a notable difference in the results between 2-positive and 1-positive FIT groups shown in our study suggests a common trend irrespective of kits brand.

We added the following section to the end of the Results.

Patients with positive FIT overlapping symptoms or history of colorectal lesions

Because FIT was conducted annually as part of colorectal cancer screening system, independent of symptoms or history of colorectal lesions, the FIT groups included patients with accompanying symptoms or history of polypectomy. In the positive FIT groups, 31 patients were symptomatic and 19 had a history of colorectal lesions. In situ cancers were found in three patients with 2-positive FIT results and haematochezia. No cancer was detected in patients with positive FIT results and history of colorectal lesions.

For the FIT (+) positive patients the authors do not indicate how many people did two tests with one positive and one negative and how many people just did the one positive test. It is possible that this latter group should have been in the FIT (2+) group if they had done the second test.

According to the reviewer's comment, we conducted a re-analysis with only patients with 1-positive results in the 2-sample FIT as the control group. The same trends were noted, and the difference seemed to be clearer. As the reviewer pointed out, the excluded groups included patients who could be classified into the FIT (+/+) group. The findings of patients with 1-positive results in the 1-sample method and with unknown number of positivity are presented in the modified Supplementary Table 1.

We added the following sentence to **Study design** of the Methods section. (Page7, Line15 in the original manuscript)

Patients with a 1-positive result for the 1-sample FIT and positive FIT results with unknown number of positivity were excluded from this study; these findings are summarised in **Supplementary Table 1**.

We changed the second sentence of the Results section (Page10, Line16-17 in the original manuscript).

From

Of these, 174 patients were excluded because they underwent colonoscopy for treatment.

To:

Of them, we excluded 174 patients for undergoing colonoscopy for treatment, 136 patients for a 1-positive result for the 1-sample FIT, and 287 patients for positive FIT results with unknown number of positivity.

In the discussion the authors suggest that FIT is better at detecting the more advanced cancers rather than the earlier lesions. I think that there is eveidence to contradict this which they have failed to discuss (Moss et al 2017 and Clark et al 2020 below)

[Ref B] *Moss S., Matthews C., et al GUT 2017:66:1631-44 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.*

[Ref C] *Clark G., Strachen JA., et al GUT 2020 doi 10.1136/gutjnl-2019-320297 . Transition to quantitative faecal immunochemical resting from guaiac faecal occult blood testing in a fully rolled-out population based national bowel screening programme.*

The sentence "the sensitivity is lower for early stage cancer or high-grade dysplasia" in the second paragraph of the Discussion section (Page 13, Line18–19 in the original manuscript, Page15, Line1-2 in the revised manuscript) might be misleding. We meant that the sensitivity of FIT considerably decreases compared with that of direct colonoscopy. FIT detects earlier lesions more effectively than the guaiac-based FOBT. This is true for the 2-sample FIT wherein patients with positive results (1- or 2-positive results) are evaluated. However, when separately analysed, 2-positive results showed extremely high positive predictive values for advanced-stage cancer, whereas 1-positive results mainly detected early-stage lesions.

We changed that sentence (Page13, Line18-19 in the original manuscript, Page15, Line1-2 in the revised manuscript) and additionally cited the references above.

From:

Although the sensitivity of FIT is superior to that of the guaiac test,[3,4] the sensitivity is lower for early stage cancer or high-grade dysplasia.

To:

Although the sensitivity of FIT is superior to that of the guaiac test,[3,4,Ref B,Ref C] it decreases considerably for early-stage cancer or high-grade dysplasia compared with direct colonoscopy.

Reviewer: 2
Dr. Marc Bardou, Centre Hospitalier Universitaire de Dijon Comments to the Author:
Revue on MS bmjopen-2020-046055 by Osomu Tosyoshimada and colleagues.
In this paper authors assessed priority stratification for colonoscopy based on the fact that patients had one or two positive FIT.
They found that Detection rates of all (and invasive) cancers in the FIT (2+), FIT (+), and non-FIT groups were 12.1% (8.3%), 1.9% (0.5%), and 0.4% (0.2%), respectively.
Although the paper may be of interest I have some significant concerns with it in its present form.

The first concern I have is that the clinical purpose of the paper is not obvious to me.
I wonder if authors want to suggest that CRC screening through the combination of 2 FIT has to become the standard approach in all countries with a FIT-based organized CRC screening program (CRCSP), or that in the case of limited access to endoscopy facilities, those with two FIT+ should be prioritized versus those with only one.
Many countries have CRCSP that are based on a single FIT result, and I don't think the paper is likely to make them change.
Here authors suggest that 2 FIT+ is more specific than 1 FIT+, whereas a CRCSP has to balance sensitivity and specificity.

The 2-sample FIT was adopted in Japan since the balance between sensitivity and specificity was considered better maintained in 2-sample methods than in 1- or 3-sample methods. The 1-sample method using quantitative FIT adjusting threshold can be simple, inexpensive, and convenient. Careful and wide-range evaluations are necessary to decide whether the 2-sample FIT is better than the 1-sampl FIT, which should depend on the various conditions of the population.

The advantage of the 2-sample FIT is based on the considerable discordance in FIT results between samples collected even from the same person. One 2-sample quantitative FIT result from another institute changed from 1 ng/mL to 1000 ng/mL (cutoff: 100 ng/mL = 20 μg haemoglobin/g faeces) on the next day. Of the 2972 patients with positive results (1- or 2-positive results from 2 samples), the discrepancy in FIT measurement values was >100 ng/mL in 87% and >200 ng/mL in 52% results (unpublished data). The 2-sample FIT may have advantages over the 1-sample FIT under some circumstances, even after adjusting the threshold. On the other hand, for risk stratification, the appropriate secondary cutoff values for the 1-sample quantitative FIT need to be decided for each FIT kit. The 2-sample FIT, using the established threshold in each FIT kit, has two possible results: 2-positive or 1-positive result.

We propose that patients with 2-positive results should be prioritised for colonoscopy, especially when resources are limited. In addition, given the COVID-19 pandemic, patients are likely to hesitate to undergo colonoscopy. In such cases, they should be strongly encouraged to receive colonoscopy with high priority. It may be useful to stratify patients with symptoms in a primary care setting, although our study cannot confirm whether the 2-sample FIT is superior to the 1-sample quantitative FIT. In the setting of 1-sample FIT screening, our results suggest that secondary FIT administered to patients with a positive primary FIT result can help identify patients at higher risk for whom colonoscopy should not be delayed.

The policy of recruiting only patients with 2-positive results has not resulted in a good balance in sensitivity and specificity for colorectal cancer screening system in Japan. Patients with 1- or 2-positive results are instructed to undergo colonoscopy. However, when resources for colonoscopy are limited, a change in the preferable balance might be possible.

In another situation, when the prevalence of target lesion is very low, for example, in the screening of younger populations, higher specificity is preferable to reduce negative findings. Although further evaluations are widely needed, the policy of recruiting only patients with 2-positive results may be worth considering for screening young generations.

We added the following paragraphs to the Discussion section, just before **Limitations and Strengths**.

The present study cannot answer whether the 2-sample FIT is superior to the 1-sample quantitative FIT as a tool for organised colorectal cancer screening program. The 1-sample FIT is simpler and less expensive at the primary screening step. Careful and wide-range evaluations are necessary to select the best method, which should depend on the various conditions of the population. An advantage of the 2-sample FIT is based on the considerable discordance in FIT results between samples collected even from the same person. The result can sometimes change from 1 ng/mL to 1000 ng/mL (cutoff:

100 ng/mL = 20 μg Hb/g faeces) by the next day. The 2-sample FIT may have advantages over the 1-sample FIT, even after adjusting the threshold, under some circumstances. On the other hand, for risk stratification, the appropriate secondary cutoff values for the 1-sample quantitative FIT need to be decided for each FIT kit. The 2-sample FIT, using the established threshold for each FIT kit, has two possible results: 2-positive or 1-positive result.

We propose that patients with 2-positive results should be prioritised for colonoscopy, especially when resources are limited. In addition, given the COVID-19 pandemic, patients are likely to hesitate to undergo colonoscopy. In such cases, they should be strongly encouraged to receive colonoscopy with high priority. It may be useful to stratify patients with symptoms in a primary care setting. In the setting of 1-sample FIT screening, our results suggest that secondary FIT administered to patients with a positive primary FIT result can help identify patients at higher risk for whom colonoscopy should not be delayed.

To be more specific.
I first have a concern with the FIT conducted in patients bellow the age of 50, but it may be because of lack of knowledge of the way the Japanese CRCSP is organised. Indeed in many countries CRC screening only starts at 50. It seems to me, from my reading of table 1, that some patients included on this study had a FIT even bellow the age of 40. Can authors comment on that?

In Japan, official colorectal screening starts at the age of 40 and is offered by the local government.

In addition, there are programs offered by employers as well as private screening programs for individuals. These screening programs are frequently incorporated into systemic medical check-ups, and age limits or screening methods are flexible in many cases. The younger population is likely to prefer programs of the latter type because of convenience. These young patients were recruited via these programs.

I think authors should rerun their analysis only including those with 2 FIT results and comparing FIT+/+ vs FIT+/- (FIT+/- being for +/- or -/+). Mixing FIT+ (i.e. those who had only one test which was positive) and FIT+/- to make a single group of one single FIT is way too confusing, particularly when haemoglobin level is not provided.

According to the reviewer's comment, we conducted a re-analysis comparing the FIT (+/+) and FIT (+/-) groups.

The same trends were noted, and the difference seemed to be clearer.

The details of patients with 1-positive result in the 1-sample FIT and unknown number of positivity are presented in the modified Supplementary Table 1.

We added the following sentence to **Study design** of the Methods section. (Page7, Line15 in the original manuscript)

Patients with a 1-positive result for the 1-sample FIT and positive FIT results with unknown number of positivity were excluded from this study; these findings are summarised in **Supplementary Table 1**.

We changed the second sentence of the Results section (Page10, Line16-17 in the original manuscript).

From

Of these, 174 patients were excluded because they underwent colonoscopy for treatment.

To:

Of them, we excluded 174 patients for undergoing colonoscopy for treatment, 136 patients for a 1-positive result for the 1-sample FIT, and 287 patients for positive FIT results with unknown number of positivity.

It would also have been useful to run the analysis based on Hb threshold, because it is well known that the lower it is, the highest is the likelihood of diagnosing non-neoplastic lesion such as adenomas.
In this extent, even if authors are unable to do this because they have no idea of HB thresholds that have been used, which is a major pitfall, it would be useful to be provided with data on non-neoplastic findings.

According to the reviewer's comment, we added information on adenomas.

We calculated the adenoma detection rate based on FIT positivity in this study population. Our result showed that the adenoma detection rate in the FIT (+/+) group was significantly higher than that in the FIT (+/-) group (61.4% vs. 47.7%, *P*<.001, using chi-squared test).

The difference in the detection rates of adenomas between the FIT (+/+) group and the FIT (+/-) group (61.4% vs. 47.7%) was less remarkable than those of invasive cancers (8.3% vs. 0.3%).

We added the following sentences to the last of **Cancer detection rates based on the indication for colonoscopy** of the Results section (Page12).

The detection rate of benign adenomas was significantly higher in the FIT (+/+) group than in the FIT (+/-) group (61.4% vs. 47.7%, P<.001). The difference in the detection rates of adenomas between the FIT (+/+) group and the FIT (+/-) group was less remarkable than those of cancers.

As it is to have the proportion of patients who were FIT+/+ and FIT+/-, and FIT-/- among all those who had two FIT results, stratified on age.
I don't think the comparison with the group of patients who underwent colonoscopy is valid, at least the way it has been done, as it is a mix of different level of risk, mostly for the surveillance group which can be really heterogeneous.

As the reviewer pointed out, positive predictive values are highly associated with the expected prevalence of lesions in the study population. Our results are susceptible to bias due to heterogeneity among our patients, which is a limitation of our study design. However, based on our results, detection rates of more advanced tumours were excellent in patients with 2-positive results, whereas they were generally quite low in the other positive groups. Further, this trend was observed irrespective of age groups. Although results could change according to the study population, we assume that higher risk for advanced-stage lesions in 2-positive FIT results is generally true for various populations.

We added the following sentences to the end of the **Limitations and Strengths**.

Fourth, positive predictive values are highly associated with the expected prevalence of lesions in the study population. Our results are susceptible to bias due to heterogeneity among our patients, which is a limitation of our study design. However, based on our results, detection rates of more advanced tumours were excellent in patients with 2-positive results, whereas they were generally quite low in the other positive groups. Further, this trend was observed irrespective of age groups. Although the results could change according to the study population, we assume that higher risk for advanced-stage lesions in 2-positive FIT results is generally true for various populations.

Other concerns
The "Ethics paragraph sound odd"
Authors write: "This study was approved by the Certificated Review Board, Hattori Clinic on September 6th, 2019 (approval no. S1909-U06, registration no. UMIN000018541). Written informed consent was obtained from the patients. All clinical investigations were conducted according to the ethical guidelines of the Declaration of Helsinki".
But to my understanding this study consists on the retrospective analysis of data collected as part of an organised CRCSP.

Thank you for your comment. This study was conducted in accordance with ethical guidelines for medical studies in Japan. Written informed consent was obtained from the patients at the time of the colonoscopy to use their data for research purpose. The study design was described in a protocol prepared by Toyoshima Endoscopy Clinic and was approved by the Certificated Review Board, Hattori Clinic on 6 September 2019 (approval no. S1909-U06, registration no. UMIN000018541). We published this study's protocol on our institute's website (http://www.ichou.com), so that patients can opt out of the study.

We modified the **Ethics** of the Methods as following.

This study was conducted in accordance with ethical guidelines for medical studies in Japan. Written informed consent was obtained from patients at the time of colonoscopy to use their data for research purposes. The study design was described in a protocol prepared by Toyoshima Endoscopy Clinic and was approved by the Certificated Review Board, Hattori Clinic on 6 September 2019

(approval no. S1909-U06, registration no. UMIN000018541). We published this study's protocol on our institute's website (http://www.ichou.com), so that patients can opt out of the study. All clinicl investigations were conducted according to the ethical guidelines of the Declaration of Helsinki.

Same for the patient and public involvement paragraph which states that "Patients and/or the public were involved in the design, or conduct, or report, or dissemination plans of this research." How is that so?

We apologize for this error. According to the reviewer's comment, we have changed this section.
From:
Patients and/or the public were involved in the design, or conduct, or report, or dissemination plans of this research. Information on the publication of this study will be provided to the patients on the website of our clinic (https://www.ichou.com).
To:
Patients and/or the public were not involved in the design, conduct, report, or dissemination of this research.

The discussion goes to fast on the limitations of the present study.

According to the reviewer's first comment, we have inserted two additional paragraphs in the Discussion just before the limitations.
According to the reviewer's comment, we have revised the **Limitations and Strengths** section as written above.

VERSION 2 – REVIEW

REVIEWER	Smith, Steve University Hospitals of Coventry & Warwickshire NHS Trust, Midlands & NW Bowel Cancer Screening Hub, Hospital of St. Cross
REVIEW RETURNED	14-Apr-2021

GENERAL COMMENTS	i am pleased to nte that the authors have made the results much clearer to understand and in doing so have addressed some of my previous comments. They have made it clear why they are unable to indicate the FIT methods used and associated threshold values with the exception of their own institution. This I feel ispotentially a significant limitation. However even with that limitation the 2 positive vs 1 positive results do give very different results and are thus applicable to any screening programme where there is no obligation to use a single type of FIT method. I am prepared to accept the authors ascertion that ."...our study suggests a common trend irrespective of brand." Because of the limitations of this work I am not convinced that this paper contributes significantly to what we already know about FIT. Page 17 Line 3 (cut off : 100ng/ml = 20ug/g) It should be made clearer that the information in the bracket applies to the OC Sensor method and is really demonstrating the conversion from one set of units.

REVIEWER	Bardou, Marc Centre Hospitalier Universitaire de Dijon, CIC-P INSERM 1432
REVIEW RETURNED	14-Apr-2021

GENERAL COMMENTS	Authors have modified their paper. I have nevertheless some remaining concerns. There are many changes questioning robustness of what have been done. For example, numbers of eligible and included patients are not

	the same between original and revised version of the paper (for example 1705 and 1282 colonoscopies respectively for FIT positivity). But at the same time some figures have not been changed, for example the number of FIT+/+ patients. The issue on positivity threshold has not been solved. Indeed, in the method section authors write that FIT were performed in their institution with a 32 mg/g threshold and in the discussion that “cutoff values for positivity were various and unknown in many cases that were referred from other medical institutions for colonoscopy” I still do think that results of the present study have not been contextualized to countries , and authors only suggest that people with 2 FIT+ should be prioritized for colonoscopy. In its way results of the present study are of no interest for countries where CRC screening is based on a single FIT test, as it does not discuss relevance of the two approaches

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Authors have modified their paper.
I have nevertheless some remaining concerns.
There are many changes questioning robustness of what have been done. For example, numbers of eligible and included patients are not the same between original and revised version of the paper (for example 1705 and 1282 colonoscopies respectively for FIT positivity).
But at the same time some figures have not been changed, for example the number of FIT+/+ patients.

We thank the reviewer for pointing out our mistake. We found that, in Table 1, the cells labelled “Age” and “Male” referring to the FIT (+/-) group were not revised. We corrected these numbers from “50.3 (12.0)” and “653 (45.3%)” to “50.3 (11.9)” and “469 (46.1%)”. We did not find any other errors.
The FIT (+/+) group remained the same as the FIT (2+) group in the original manuscript because we excluded the FIT (+, in 1 sample) and FIT (+/unknown result) groups from the mixed FIT (+) group, and the population of the FIT (+/-) group was reduced from the original FIT (+) group.

The issue on positivity threshold has not been solved. Indeed, in the method section authors write that FIT were performed in their institution with a 32 mg/g threshold and in the discussion that “cutoff values for positivity were various and unknown in many cases that were referred from other medical institutions for colonoscopy”
I still do think that results of the present study have not been contextualized to countries , and authors only suggest that people with 2 FIT+ should be prioritized for colonoscopy.

As the reviewer pointed out, our main conclusion is “people with 2 FIT+ should be prioritized for colonoscopy.” For risk stratification, 2 FIT+ could identify efficiently, specifically for patients with advanced tumours whose colonoscopy should not be delayed. The judgement criterion is simple, without the need to decide the additional cutoff values and probably effective irrespective of FIT kits brand.

In its way results of the present study are of no interest for countries where CRC screening is based on a single FIT test, as it does not discuss relevance of the two approaches

As discussed in the manuscript, we don't claim that the 1-sample quantitative FIT should be replaced by the 2-sample FIT. We proposed that 2-sample FIT could be helpful when colonoscopy resources are limited, or when people are likely to avoid colonoscopy. In the countries or targets without standard screening methods, such as younger populations, 2-sample FIT may be worth considering. We hope our results can be useful in some countries or some situations.

Reviewer: 1
i am pleased to nte that the authors have made the results much clearer to understand and in doing so have addressed some of my previous comments. They have made it clear why they are unable to indicate the FIT methods used and associated threshold values with the exception of their own institution. This I feel ispotentially a significant limitation. However even with that limitation the 2 positive vs 1 positive results do give very different results and are thus applicable to any screening programme where there is no obligation to use a single type of FIT method. I am prepared to accept the authors ascertion that ."...our study suggests a common trend irrespective of brand."
Because of the limitations of this work I am not convinced that this paper contributes significantly to what we already know about FIT.
Page 17 Line 3 (cut off : 100ng/ml = 20ug/g) It should be made clearer that the information in the bracket applies to the OC Sensor method and is really demonstrating the conversion from one set of units.

Thank you very much for your insightful comments, which have considerably improved our manuscript. We changed the sentence mentioned as follows.
From: The result can sometimes change from 1 ng/mL to 1000 ng/mL (cutoff: 100 ng/mL = 20 µg Hb/g faeces) by the next day.
To: The result can sometimes change from 1 ng/mL for the first sample to 1000 ng/mL for the second sample on the next day (cutoff: 100 ng/mL = 20 µg Hb/g faeces, in the case of the OC Sensor method, Eiken Chemical Co., Ltd., Tokyo, Japan).

VERSION 3 – REVIEW

REVIEWER	Bardou, Marc Centre Hospitalier Universitaire de Dijon, CIC-P INSERM 1432
REVIEW RETURNED	28-Apr-2021

GENERAL COMMENTS	No further comments to make.