PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Comparative Evaluation of Novel Screening Strategies for Colorectal Cancer Screening in China (TARGET-C): A Study Protocol for a Multicenter Randomized Controlled Trial
AUTHORS	Chen, Hongda; Li, Ni; Shi, Jufang; Ren, Jian-song; Liu, Chengcheng; Zhang, Yueming; Jiang, Zheng; Zhang, Zhihui; Dai, Min

VERSION 1 - REVIEW

REVIEWER	Peter S. Liang New York University Langone Health / VA New York Harbor Health Care System, USA
REVIEW RETURNED	06-Sep-2018

GENERAL COMMENTS	Summary This is the first randomized controlled trial of colonoscopy vs. FIT outside of the Western Hemisphere, and as such the results will be of great significance to the field of colorectal cancer (CRC) screening. In addition, the investigators have also added a third arm of risk-stratifying participants into FIT or colonoscopy according to the Asia-Pacific Colorectal Cancer score. This is a novel intervention that is not currently being studied in any of the other ongoing CRC screening trials. The authors should be commended for what promises to be a highly impactful study. Nevertheless, the protocol could be strengthened by a number of clarifications. Major comments 1. The choice to use advanced adenoma detected as the primary outcome, rather than CRC incidence and mortality, is not adequately explained. The authors only mention that they may not have adequate sample size to compare CRC outcomes, but these are the most clinically relevant and important outcomes. Thus, regardless of whether the advanced adenoma detection is different in the different arms, most readers would be most interested to see the CRC detection rates. 2. The authors seem most interested to demonstrate the non-inferiority of colonoscopy vs. APCS score for advanced neoplasia detection rate. If their hypothesis is correct, it seems that the argument would be using the APCS score is most cost-effective. However, they state the advanced neoplasia detection rate is 6.5% for colonoscopy and 5.0% for APCS. This implies that APCS misses 1.5/6.5=23% of advanced neoplasias that are detected by

	colonoscopy. It seems that the only way there will be non-inferiority is if there is a large discrepancy in compliance. 3. The test performance characteristics (e.g., sensitivity, specificity, PPV) of the particular qualitative FIT that is being used in this study and the APCS should be provided. 4. The protocol would benefit from proofreading for language and grammar. Minor comments Abstract - Methods and analysis: the sample size is written as 200,000 Introduction -The first sentence is contradictory. The authors presumably wanted to refer to cancer death. -2nd paragraph: I disagree that the fecal DNA test is widely used. -2nd paragraph->last sentence: should be “still lacking”. Methods -Study design: the description of 3 vs. 10 years of follow-up is confusing. The authors should explain “active” vs. “passive” follow-up earlier. -FIT group: participants with positive FIT should undergo diagnostic, not screening, colonoscopy. -It seems like only participants who undergo colonoscopy receive a annual questionnaire. It’s unclear why the FIT patients would not receive the questionnaire. -Are participants with a prior history of adenomas not excluded? If so, then this is a mixed rather than screening population. -Is the upper limit of age eligibility 74 or 75? Both number are used in the protocol. -Why is testing for HBV, HCV, and HIV performed prior to colonoscopy? If this is standard practice in China, this should be explained. -Are self-reported FIT’s reliable? -Colorectal cancer risk assessment: BMI is not listed as one of the five risk factors. -Follow up: For active follow up, it’s unclear what “diagnostic examination” means. Is it a physical examination? -Sample size: I would advise against using the abbreviation ADR to refer to "advanced neoplasia detection rate," because that refers to the adenoma detection rate by convention. Discussion -Correction: the NordICC trial compares colonoscopy vs. no screening. There is no FIT arm. Tables/Figures -Table 1: title needs to be re-formatted SPIRIT checklist: -There is no mentioning of a data monitoring committee or the logging of adverse events (e.g., perforation, bleeding)

REVIEWER	Isabel Idigoras Rubio Osakidetza, Basque Health Service, Spain
REVIEW RETURNED	08-Oct-2018

GENERAL COMMENTS	Confusion of widely used terminology such as ADR = adenoma detection rate instead of Advanced adenoma rate FIT Qualitative, with lower sensitivity and specificity than quantitative Even without results there should be better explanation about the 3 arms of the study

VERSION 1 – AUTHOR RESPONSE

Reviewers' Comments to Author:

Reviewer: 1

Reviewer Name: Peter S. Liang

Institution and Country: New York University Langone Health / VA New York Harbor Health Care System, USA

Please state any competing interests or state ‘None declared’: None declared

Summary This is the first randomized controlled trial of colonoscopy vs. FIT outside of the Western Hemisphere, and as such the results will be of great significance to the field of colorectal cancer (CRC) screening. In addition, the investigators have also added a third arm of risk-stratifying participants into FIT or colonoscopy according to the Asia-Pacific Colorectal Cancer score. This is a novel intervention that is not currently being studied in any of the other ongoing CRC screening trials. The authors should be commended for what promises to be a highly impactful study. Nevertheless, the protocol could be strengthened by a number of clarifications.

Major comments

1. The choice to use advanced adenoma detected as the primary outcome, rather than CRC incidence and mortality, is not adequately explained. The authors only mention that they may not have adequate sample size to compare CRC outcomes, but these are the most clinically relevant and important outcomes. Thus, regardless of whether the advanced adenoma detection is different in the different arms, most readers would be most interested to see the CRC detection rates.

 Response: Thank you for your comments. We have made the changes in the manuscript to make the primary and secondary outcomes identical between the study protocol and the registry. The primary outcome is the colorectal cancer mortality rate. The secondary outcomes include detection rate of CRC, detection rate of precancerous lesions of CRC, compliance rate, complication rate (page 16, lines 1-6).

2. The authors seem most interested to demonstrate the non-inferiority of colonoscopy vs. APCS score for advanced neoplasia detection rate. If their hypothesis is correct, it seems that the argument would be using the APCS score is most cost-effective. However, they state the advanced neoplasia detection rate is 6.5% for colonoscopy and 5.0% for APCS. This implies that APCS misses 1.5/6.5=23% of advanced neoplasias that are detected by colonoscopy. It seems that the only way there will be non-inferiority is if there is a large discrepancy in compliance.

 Response: Thank you for your comments. We indeed assumed that the risk-adapted approach is non-inferior to the colonoscopy approach given the potential higher compliance rate of risk-adapted approach than traditional colonoscopy. Previous studies have shown that the overall participation rate of colonoscopy was poor in China (Gut 2018.doi: 10.1136/gutjnl-2018-317124), and FIT has a much

better participation rate (Gut. 2017;66(9):1631-1644.). Given this we anticipate that risk-adapted approach is superior to the FIT approach and non-inferior to the colonoscopy approach.

3. The test performance characteristics (e.g., sensitivity, specificity, PPV) of the particular qualitative FIT that is being used in this study and the APCS should be provided.

 Response: Thank you for your comments. We have tested the diagnostic performance of the FIT in a pilot study, the overall diagnostic accuracy was satisfactory. But the data is not public available, therefore we decided not to provide such information in the main text. The test performance of APCS has been added (page 13, lines 3-5).

4. The protocol would benefit from proofreading for language and grammar.

 Response: Done as suggested.

Minor comments

Abstract

- Methods and analysis: the sample size is written as 200,000

 Response: Corrected as suggested.

Introduction

-The first sentence is contradictory. The authors presumably wanted to refer to cancer death.

 Response: Corrected as suggested (page 5, line 3).

-2nd paragraph: I disagree that the fecal DNA test is widely used.

 Response: Thank you for your comments. We have deleted fecal DNA tests in the sentence (page 5, line 12).

-2nd paragraph->last sentence: should be “still lacking”.

 Response: Corrected as suggested (page 6, line 1).

Methods

-Study design: the description of 3 vs. 10 years of follow-up is confusing. The authors should explain “active” vs. “passive” follow-up earlier.

 Response: Thank you for your comments. We have rephrased the sentence to make address this issue (page 8, lines 9-13).

-FIT group: participants with positive FIT should undergo diagnostic, not screening, colonoscopy.

 Response: Done as suggested (page 9, line 4).

-It seems like only participants who undergo colonoscopy receive a annual questionnaire. It’s unclear why the FIT patients would not receive the questionnaire.

 Response: A standardized epidemiological questionnaire survey will be conducted for all participants at baseline screening. For the next three rounds of screening, because the colonoscopy group will not receive any further intervention, we therefore plan to have a questionnaire survey to obtain the health status of the participants each year.

-Are participants with a prior history of adenomas not excluded? If so, then this is a mixed rather than screening population.

 Response: Thank you for your comments. We don't exclude participants have a prior history of adenoma. Such information will be collected in the epidemiological questionnaire survey, and will be taken into consideration in the future analyses.

-Is the upper limit of age eligibility 74 or 75? Both number are used in the protocol.

 Response: The upper limit of age eligibility is 74. We have corrected errors in the main text.

-Why is testing for HBV, HCV, and HIV performed prior to colonoscopy? If this is standard practice in China, this should be explained.

 Response: Blood tests for infectious diseases (such as HBV, HCV and HIV) were required in some hospitals in China but not a standard procedure by all hospitals. We have added a short explanation about this (page 11, lines 14-15).

-Are self-reported FIT's reliable?

 Response: Yes, the self-reported FIT is reliable, which have been also used in our pilot study (data has been published). As state in the method section, we have developed several approaches to enhance the accuracy of the self-reported results, including submitting the results along with the pictures of test window, and auxiliary interpretation of the results by trained staff.

-Colorectal cancer risk assessment: BMI is not listed as one of the five risk factors.

 Response: Corrected as suggested (page 13, line 3).

-Follow up: For active follow up, it's unclear what "diagnostic examination" means. Is it a physical examination?

 Response: Thank you for your comments. We have corrected it to" physical examination" (page 15, line 6).

-Sample size: I would advise against using the abbreviation ADR to refer to "advanced neoplasia detection rate," because that refers to the adenoma detection rate by convention.

 Response: Thank you for your comments. The abbreviation of ADR has been abandoned in the main text to avoid mindreading.

Discussion

-Correction: the NordICC trial compares colonoscopy vs. no screening. There is no FIT arm.

 Response: Thank you for your comments. We have corrected this error (page 20, line 9).

Tables/Figures

-Table 1: title needs to be re-formatted

 Response: Thank you for your comments. Now we have re-formatted the title of table 1 (page 27).

SPIRIT checklist:

-There is no mentioning of a data monitoring committee or the logging of adverse events (e.g., perforation, bleeding)

 Response: Thank you for your comments. We had added such information in the main text (page 16, lines 21-22; page 18, lines 1-7).

Reviewer: 2

Reviewer Name: Isabel Idigoras Rubio

Institution and Country: Osakidetza, Basque Health Service, Spain

Please state any competing interests or state ‘None declared’: No conflicts of interest

Confusion of widely used terminology such as ADR = adenoma detection rate instead of Advanced adenoma rate

 Response: Thank you for your comments. The abbreviation of ADR has been abandoned in the main text to avoid mindreading.

FIT Qualitative, with lower sensitivity and specificity than quantitative

 Response: In large-scale population-based cancer screening programs, feasibility and operability of study protocol must be taken into consideration. Although the quantitative FIT has several advantages over qualitative FIT, quantitative FIT typically needs recollection, centrally processing and testing the samples. Such processes may affect the uptake rate of the participants. In our previous pilot study, the qualitative FIT used in this trial exhibited overall good performance. Of note, the participants could performance the test at home easily and therefore strongly improved the overall participation rate. As state in the method section, we have developed several approaches to enhance the accuracy of the self-reported results, including submitting the results along with the pictures of test window, and auxiliary interpretation of the results by trained staff. We are confident with the overall performance of this qualitative FIT.

Even without results there should be better explanation about the 3 arms of the study

 Response: Thank you for your comments. Detailed information about the designs the 3 arms have been detailed described in the method section. Colonoscopy is the gold standard for colonoscopy screening and is therefore used to be as reference. FIT is the most widely used non-invasive test for colonoscopy screening test. In countries having relatively low incidence of CRC or having limited health resources, a risk score using to identify high-risk population for screening is recommended by the consensus in Asia and also in China. We therefore developed a novel risk-adapted screening strategy in our trial. And the overall diagnostic performance of this risk-adapted screening strategy was reported in previous studies (Gut, 2011. 60(9): 1236-41; Gastroenterology, 2016. 150(3): 617-25.e3). Successful implementation of this study will provide strong evidence on designing suitable strategies for CRC screening in China and other countries.

VERSION 2 – REVIEW

REVIEWER	Peter S. Liang NYU Langone Health and VA New York Harbor Health Care System USA
REVIEW RETURNED	29-Dec-2018

GENERAL COMMENTS	The manuscript has been substantially improved in this revision, and the authors should be commended for their efforts. Nevertheless, two points from the original review have still not been adequately addressed. 1. The authors state they have made CRC mortality the primary outcome. While this has been changed in the "Outcome measures" section, it has not been changed in the Abstract, "Statistical Considerations," or "Statistical analyses," where advanced neoplasia detection rate is still the primary outcome. Notably, the sample size calculations are still for advanced neoplasia detection, not CRC mortality. 2. The authors have declined to provide the test performance characteristics of the qualitative FIT used in the trial, but this seems to be an important piece of information. In their sample size calculation, they state the advanced neoplasia detection rate for FIT as 1.8%, but the references provided are for a quantitative FIT and not the one being used in the trial. Also, for the APCS score they provided the relative prevalence of advanced neoplasm (2.48-fold) in the high vs. low-risk participants, but this is not the same or as useful as the sensitivity, specificity, and PPV of the score.

VERSION 2 – AUTHOR RESPONSE

Point-by-point responses to reviewer's comments

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name: Peter S. Liang

Institution and Country: NYU Langone Health and VA New York Harbor Health Care System

USA

Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below

The manuscript has been substantially improved in this revision, and the authors should be commended for their efforts. Nevertheless, two points from the original review have still not been adequately addressed.

1. The authors state they have made CRC mortality the primary outcome. While this has been changed in the "Outcome measures" section, it has not been changed in the Abstract, "Statistical Considerations," or "Statistical analyses," where advanced neoplasia detection rate is still the primary outcome. Notably, the sample size calculations are still for advanced neoplasia detection, not CRC mortality.

 Response: Thank you for your comments. After consulting with the statisticians and having a thorough internal discussion, we decided to adhere to previous primary outcome, i.e., detection rate of advanced neoplasm (including CRC and advanced adenoma), the main reason is because the

planned sample size (20,000) does not have the power to compare the 10-year mortality rate of CRC between the 3 study groups. To make it consistency, the online registry has also been modified accordingly. We agree with the reviewer that the mortality rate is quite essential for evaluating the screening strategies, we might expand the sample size if the interim results are satisfactory in the future. Please refer the changes made in page 15, lines 18-22.

2. The authors have declined to provide the test performance characteristics of the qualitative FIT used in the trial, but this seems to be an important piece of information. In their sample size calculation, they state the advanced neoplasia detection rate for FIT as 1.8%, but the references provided are for a quantitative FIT and not the one being used in the trial. Also, for the APCS score they provided the relative prevalence of advanced neoplasm (2.48-fold) in the high vs. low-risk participants, but this is not the same or as useful as the sensitivity, specificity, and PPV of the score.

 Response: Thank you for your comments. We agree with you that the characteristics of FIT is essential in evaluating the performance of the screening strategy. The diagnostic performance of qualitative FITs may vary greatly due to the different positive threshold values used by the manufacturers (Ann Intern Med. 2009 Feb 3;150(3):162-9; Gastroenterology. 2018 Jan;154(1):93-104). In the qualitative FIT used in our trial, the positive threshold is 100ng Hb/ml buffer, which corresponds to 10 μg Hb/g feces. In a previous pilot study (data not publicly available), the sensitivities for detecting CRC and advanced adenomas was 76% and 37% respectively, at a specificity of 92%, which was comparable to other FITs reported in other studies. We used the 1.8% as the reference detection rate for advanced neoplasia given the test characteristics is comparable between the FIT used in our study and the FIT (OC FIT-CHEK, Polymedco, positive threshold of 100 ng Hb/ml buffer) used in the study by Imperial et al. As we lacked validation of the test performance of the qualitative FIT in large-scale screening populations, we have to use previous research results using other FIT brand having the similar test characteristics as a compromised way to address this important issue. Sensitivity analysis changing the detection rate slightly (±0.1%) still indicates the current sample size meet the research hypotheses. For the APCS score, the diagnostic indicators have been provided as suggested (page 12, lines 18-21; page 13, lines 1-6).

VERSION 3 - REVIEW

REVIEWER	Peter S. Liang NYU Langone Health VA New York Harbor Health Care System
REVIEW RETURNED	11-Jan-2019

GENERAL COMMENTS	The authors have adequately addressed my previous comments. I have no additional concerns.

VERSION 3 – AUTHOR RESPONSE

Reviewer's Comments to Author:

Reviewer: 1

Reviewer Name: Peter S. Liang

Institution: NYU Langone Health, VA New York Harbor Health Care System

Please state any competing interests or state ‘None declared’: None declared

The authors have adequately addressed my previous comments. I have no additional concerns.

Responses: Thank you for your comments.