Faecal immunochemical tests versus colonoscopy for post-polypectomy surveillance: an accuracy, acceptability and economic study

Wendy Atkin,[1*†] Amanda J Cross,[1*] Ines Kralj-Hans,[1] Eilidh MacRae,[1] Carolyn Piggott,[2] Sheena Pearson,[2] Kate Wooldrage,[1] Jeremy Brown,[1] Fiona Lucas,[1] Aaron Prendergast,[1] Natalie Marchevsky,[1] Bhavita Patel,[1] Kevin Pack,[1] Rosemary Howe,[1] Hanna Skrobanski,[3] Robert Kerrison,[3] Nicholas Swart,[4] Julia Snowball,[2] Stephen W Duffy,[5] Stephen Morris,[4] Christian von Wagner[3] and Stephen Halloran[2]

[1]Cancer Screening and Prevention Research Group, Department of Surgery and Cancer, Imperial College London, London, UK
[2]Bowel Cancer Screening Programme Southern Hub, Guildford, UK
[3]Research Department of Behavioural Science and Health, University College London, London, UK
[4]Department of Applied Health Research, University College London, London, UK
[5]Centre for Cancer Prevention, Wolfson Institute of Preventative Medicine, Queen Mary University, London, UK

*Corresponding authors amanda.cross@imperial.ac.uk
†In memoriam

Declared competing interests of authors: Wendy Atkin and Amanda J Cross report grants from the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) programme, grants from Cancer Research UK (Population Research Committee – Programme Award C8171/A16894) and non-financial support from Eiken Chemical Co. Ltd (Tokyo, Japan) (MAST is UK distributor) during the conduct of the study. Stephen Morris is a member of the NIHR Health Services and Delivery Research funding board. Sheena Pearson, Carolyn Piggott and Julia Snowball all report grants from the NIHR HTA programme during the conduct of the study.

Disclaimer: This is a summary of independent research funded by the NIHR HTA programme and the Bobby Moore Fund for Cancer Research UK. The views expressed are those of the authors and not necessarily those of the NHS, NIHR, the Department of Health and Social Care or Cancer Research UK. Infractructure support for this work was provided by the NIHR Imperial Biomedical Research Centre.

Published January 2019
DOI: 10.3310/hta23010

Scientific summary

Post-polypectomy surveillance: FIT vs. colonoscopy

Health Technology Assessment 2019; Vol. 23: No. 1
DOI: 10.3310/hta23010

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Colorectal cancer (CRC) causes considerable morbidity and mortality in the UK and internationally. Most CRCs develop from precursor lesions called adenomas. Adenoma removal through polypectomy reduces CRC incidence; however, even after adenomas have been removed, many patients remain at increased risk of CRC.

Given the increased risk of CRC in patients post polypectomy, national guidelines in the USA, the UK, the European Union and elsewhere recommend surveillance of these patients at regular intervals using colonoscopy. The length of the surveillance interval depends on the number and features of adenomas found, including size and histology. Following polypectomy, the UK adenoma surveillance guidelines divide patients into low-, intermediate- and high-risk groups. Low-risk patients, defined as those with one or two small adenomas (i.e. sized < 10 mm), are recommended either no surveillance or surveillance every 5 years. The Bowel Cancer Screening Programme (BCSP) in England currently adopts the no surveillance approach, with low-risk patients returning to biennial guaiac faecal occult blood test (gFOBT) screening for as long as they remain eligible. Intermediate-risk patients, namely, those with three or four small adenomas (i.e. sized (< 10 mm), or one or two adenomas with at least one sized ≥ 10 mm, are recommended to undergo surveillance every 3 years. High-risk patients, namely, those with five or more adenomas, or three or more adenomas with at least one sized ≥ 10 mm, are recommended annual surveillance. Post-polypectomy colonoscopy surveillance has been shown to reduce CRC incidence in several studies.

Although colonoscopy has a high sensitivity for CRC and advanced adenomas (AAs) (i.e. sized ≥ 10 mm, tubulovillous or villous histology, or high-grade dysplasia), it carries a small risk of complications, is time-consuming and can cause discomfort for patients. Furthermore, demand on endoscopists and the cost of colonoscopies to the NHS are increasing because of CRC screening and reductions in the referral threshold in primary care for patients with suspected CRC. The National Institute for Health and Care Excellence (NICE) guidelines now recommend referral for patients with symptoms and signs conferring a positive predictive value (PPV) for CRC as low as 3% (National Institute for Health and Care Excellence (NICE). *Suspected Cancer: Recognition and Referral.* London: NICE; 2015).

The majority of surveillance colonoscopies do not detect CRC. An alternative to colonoscopy for post-polypectomy surveillance may be the faecal immunochemical test (FIT). Like gFOBT, a FIT detects haemoglobin from blood in stool, although FITs detect species-specific globin rather than haem. Biennial gFOBT screening has been demonstrated to reduce CRC mortality in randomised controlled trials. Compared with gFOBT, FITs are less susceptible to dietary interference and more specific to lower gastrointestinal tract bleeding. Analysis of FITs can be automated, is not subject to screener interpretation and, for quantitative FIT, the positivity threshold can be modified to yield defined positivity rates. Furthermore, at low thresholds, FITs have higher sensitivity than gFOBT for CRC and AAs. Given these advantages, FIT has been adopted by many screening programmes and is set to replace gFOBT in the BCSP in England.

Although many studies have evaluated the diagnostic accuracy of FIT for CRC and AAs in screening, few have examined FIT performance in surveillance. It was hypothesised that annual FITs could be a safe, effective and cost-saving alternative to colonoscopy for surveillance of intermediate-risk patients post polypectomy.

Objectives

The primary objective was to determine the 3-year programme sensitivity of annual FITs compared with colonoscopy surveillance at 3 years for detecting CRC or AAs in intermediate-risk patients, following polypectomy after a positive gFOBT.

Secondary objectives were to:

- estimate the diagnostic accuracy of FITs at first, second and third tests and over two or three tests at various thresholds
- examine the acceptability of FITs, compared with colonoscopy, as a surveillance strategy for people at increased risk of CRC
- calculate the incremental costs and cost-effectiveness of FITs versus colonoscopy surveillance.

Methods

Eligible individuals were those aged 60–72 years who were categorised as being at intermediate risk following polypectomy at colonoscopy conducted < 1 year previously following a positive gFOBT in the BCSP, and were scheduled for surveillance colonoscopy 3 years after initial colonoscopy in line with UK guidelines [Atkin WS, Saunders BP, British Society for Gastroenterology. Surveillance guidelines after removal of colorectal adenomatous polyps. *Gut* 2002;**51**(Suppl. 5):V6–9].

Consecutive individuals meeting the eligibility criteria were sent a FIT and invited to participate in the study by the BCSP Southern Hub between January 2012 and December 2013. The FIT kit contained instructions, an information sheet, a FIT sampling device (OC-AUTO Sampling Bottle 3, Eiken Chemical Co. Ltd, Tokyo, Japan), a plastic zip-lock bag and a pre-paid envelope in which to return the completed kit and consent form.

Eligible individuals who returned a completed consent form and analysable FIT were included in the study. Laboratory analysis of FITs was conducted using the OC-Sensor DIANA (Eiken Chemical Co. Ltd, supplied by MAST Diagnostics Division, UK). The faecal haemoglobin threshold was initially 20 µg of haemoglobin/g faeces(hereafter referred to as µg/g); however, as positivity was higher than expected in the pilot study, the threshold was raised to 40 µg/g.

The study involved three rounds of FIT, conducted at 1, 2 and 3 years post baseline colonoscopy. Only participants returning a round 1 FIT were invited to subsequent rounds. Participants testing positive at round 1 or 2 were offered early colonoscopy and were not invited to further rounds. Participants testing negative at round 1 were invited to round 2. A round 3 FIT was sent to participants testing negative at round 2, and to participants who did not return an analysable round 2 FIT. Participants invited to round 3 were scheduled for colonoscopy at 3 years post baseline colonoscopy, in line with UK guidelines (Atkin and Saunders 2002).

We calculated uptake and positivity of FIT at each round. We analysed sensitivity, specificity, PPVs and negative predictive values (NPVs) of FIT for CRC, AAs and advanced colorectal neoplasia (ACN) (CRC and/or AAs) at various thresholds (40 µg/g, 30 µg/g, 20 µg/g and 10 µg/g) and over multiple rounds. We assumed that any ACN detected was present and would remain present and unchanged in the absence of colonic examination at years 1, 2 and 3, and that the same neoplasia would be detected regardless of the year at which examination occurred.

For analysis of multiple rounds, we performed a cumulative test analysis and a programme analysis. In the programme analysis, we included all participants and categorised as positive anyone testing positive in the first or second round (two-tests analysis), or in the first, second or third round (three-tests analysis). For the cumulative test analysis, we included only participants who were compliant with testing (i.e. they completed the

specified number of rounds, two for two-test analysis or three for three-test analysis, or tested positive at a previous round).

Health psychology assessment

There were two components to the health psychology assessment: (1) a qualitative discussion group study assessing attitudes towards FIT as an alternative to colonoscopy surveillance in adults with varied CRC risk and experience of colonoscopy, and (2) an evaluation of the psychological impact and acceptability of annual FIT and preferences for future surveillance using questionnaires and interviews.

For the discussion groups, 198 adults aged 60–74 years, with different levels of CRC risk and varying amounts of experience with gFOBT and colonoscopy, were identified by the BCSP London Hub and St Mark's Hospital Endoscopy Unit.

Five discussion groups were held in 2011 using a comprehensive stepwise discussion guide. After each section, participants were asked to 'consider the information you have just seen about FIT replacing a routine colonoscopy. How would you feel about the offer of a FIT every year instead of a 3-yearly colonoscopy?'. Participants used an electronic device to select an option on a six-point scale from 'very positive' to 'very negative'.

The views of FIT for Follow-Up study participants were gathered through questionnaires at baseline and at each round. In addition, a subsample of participants underwent end-of-study interviews. In the baseline questionnaire, participants were asked to rate their experience of baseline colonoscopy, level of CRC-related worry and current emotional state using the Spielberger State–Trait Anxiety Inventory (STAI). The questionnaires queried participants' experience of completing FIT. Questionnaires at each round queried emotional well-being and CRC-related worry.

Economic evaluation

We assessed incremental costs and cost-effectiveness of annual FIT with colonoscopy only for a positive result, as an alternative to colonoscopy at 3 years, for the surveillance of intermediate-risk patients.

The cost of each surveillance regimen was estimated. Cost-effectiveness is presented as the incremental cost-effectiveness (i.e. saving) per AA that was not detected by FIT versus colonoscopy surveillance, and the incremental cost-effectiveness per CRC that was not detected by FIT versus colonoscopy surveillance.

Costs were restricted to those from surveillance and did not include treatment costs, other than the cost of adenoma removal as a result of FIT positivity. Unit costs were obtained from the *NHS Reference Costs 2014 to 2015* (Department of Health and Social Care. *NHS Reference Costs 2014 to 2015*. London: Department of Health and Social Care; 2015). Sensitivity analyses explored the effect of different FIT thresholds and diagnostic procedure costs on cost-effectiveness. We estimated the budget impact if annual FIT surveillance was implemented nationally instead of 3-yearly colonoscopy surveillance.

Results

Of the 9851 individuals identified as potentially eligible, 296 were excluded and 1547 were not invited because the target sample size had been reached. The remaining 8008 were invited to participate and 5948 (74.3%) consented and completed a FIT. Two individuals subsequently withdrew, leaving a cohort of 5946 individuals in round 1. Uptake of FIT in rounds 2 and 3 was 97.2% (5350/5503) and 96.9% (5058/5220), respectively.

With a threshold of 40 µg/g, FIT positivity was 5.8% (347/5946) in round 1, decreasing to 4.1% (206/5058) in round 3. Positivity in each round was greater in men than in women (round 1: 6.6% vs. 4.3%, respectively) and in older participants (> 65 years) than in younger participants (≤ 65 years) (round 1: 6.5% vs. 5.1%,

respectively). Over all three rounds, cumulative attendance for colonic examination following a positive FIT was 93.6% (744/795). Among participants with only negative FITs, attendance for end-of-study colonic examination was 88.4% (4455/5039).

Among the 5199 participants who had a colonic examination, either following a positive FIT ($n = 744$) or at the end of the study ($n = 4455$), CRC was identified in 0.5% (26/5199) and AAs in 8.5% (443/5199), with both being identified in six participants. In total, 8.9% (463/5199) of participants were diagnosed with ACN (CRC and/or AA). Over all three rounds, 69.2% (18/26) of participants with CRC, 34.3% (152/443) with AAs and 35.6% (165/463) with ACN tested FIT positive with the 40 µg/g threshold.

FIT positivity increased at lower thresholds; for instance, 5.8% (344/5946) tested positive with the first FIT at 40 µg/g, whereas 14.2% (844/5946) tested positive at 10 µg/g. Sensitivity increased, whereas specificity decreased, with lower thresholds. Sensitivity and specificity of the first FIT for CRC were 30.8% (8/26) and 93.9% (4855/5173), respectively, at 40 µg/g, and 61.5% (16/26) and 86.0% (4447/5173), respectively, at 10 µg/g.

Sensitivity for AAs was lower, and specificity higher, than for CRC. For example, sensitivity and specificity of the first FIT for AAs were 17.6% (78/443) and 94.8% (4508/4756), respectively, at 40 µg/g, and 33.2% (147/443) and 87.5% (4161/4756), respectively, at 10 µg/g.

Taking into account multiple FIT rounds, sensitivity increased but specificity decreased. In programme analysis, sensitivity and specificity for CRC at 40 µg/g were, respectively, 61.5% (16/26) and 89.5% (4630/5173) over two rounds, and 69.2% (18/26) and 86.0% (4450/5173) over three. Similarly, in cumulative test analysis, sensitivity and specificity for CRC at 40 µg/g were, respectively, 66.7% (16/24) and 89.4% (4596/5139) for two tests, and 81.8% (18/22) and 85.4% (4243/4966) for three tests.

Health psychology assessment

In total, 28 people with different levels of CRC risk took part in five qualitative discussion groups. Previous experience of surveillance and level of CRC risk were associated with attitudes towards FIT. All groups thought that FIT would be easier and safer as a surveillance method than colonoscopy. However, individuals with prior experience of surveillance were concerned about the ability of FIT to detect lesions, and particularly that single-sample FIT might not detect polyps that bleed intermittently.

FIT for Follow-Up study participants were invited to complete questionnaires during the study, and a subset of FIT-positive participants was invited to be interviewed at the end of each round. A baseline questionnaire was completed by 98.9% (5879/5946) of participants. Questionnaires were completed by 84.4% (5020/5946) of participants at the end of round 1, 83.9% (4491/5350) at the end of round 2 and 83.5% (3881/4646) at the end of round 3.

In the baseline questionnaire, the vast majority of participants (95.8%, 5370/5604) were satisfied with their baseline colonoscopy. Most participants reported that catching the bowel motion, removing the stick, collecting the sample with the stick, reinserting the stick into the sampling bottle and closing the sampling bottle (94.9%, 99.0%, 97.1%, 95.4% and 99.3%, respectively) was easy.

Of round 1 questionnaire responders, 26.8% (1307/4877) reported that doing FIT made them anxious and 29.2% (1416/4856) reported that they were concerned about the ability of FIT to detect new polyps. Examining STAI scores, anxiety was higher at baseline than at rounds 1, 2 and 3. Participants reported high levels of reassurance, feeling more hopeful and less anxious as a result of participating in the study.

Participants' preferred surveillance strategy was 3-yearly colonoscopy plus annual FIT (57.9%, 2478/4279), followed by annual FIT with colonoscopy only for a positive result (31.5%, 1347/4279). The least preferred strategies included 3-yearly colonoscopy and no FIT (8.9%, 379/4279), and no surveillance (1.8%, 75/4279).

In end-of-study interviews, participants generally reported being very satisfied with taking part in the study. They reported that a FIT was easier to complete than a gFOBT and they appreciated the reassurance that annual FIT provided.

Economic evaluation

Among the 5946 study participants, we estimated the cost of colonoscopy surveillance at 3 years to be £2,633,382, whereas the cost of annual FIT surveillance using a threshold of 40 µg/g was £485,236. However, FITs were also less effective at detecting ACN. Using a threshold of 40 µg/g, three rounds of FITs missed 291 AAs and eight CRCs. The incremental cost-effectiveness (i.e. saving) from FITs was £7382 per AA not detected and £268,518 per CRC not detected.

The estimated cost of FITs depended on the threshold, ranging from £485,236 using 40 µg/g to £956,602 using 10 µg/g. The incremental cost-effectiveness improved with lower thresholds because of fewer missed lesions. At the lowest studied threshold, the incremental cost-effectiveness was £8872 per AA not detected and £419,195 per CRC not detected. The budget impact of replacing colonoscopy surveillance with FIT surveillance ranged from –£4.6M at a threshold of 40 µg/g to –£3.6M at a threshold of 10 µg/g.

Conclusions

This study demonstrated the potential utility of FITs in the surveillance of intermediate-risk patients post polypectomy. Annual low-threshold FIT with colonoscopy in positive cases achieved high cumulative sensitivity for CRC and would be cost saving compared with 3-yearly colonoscopy. However, at higher thresholds, this strategy could miss 15–30% of CRCs and 40–70% of AAs. Participants' preferred surveillance strategy was annual FIT plus 3-yearly colonoscopy. Further research is needed to define a clear role for FITs in surveillance, including evaluation of the implications of missed ACN, considering effects on quality-adjusted life-years.

Trial registration

This trial is registered as ISRCTN18040196.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research (NIHR), the NIHR Imperial Biomedical Research Centre and the Bobby Moore Fund for Cancer Research UK. MAST Group Ltd provided FIT kits.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.513

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 09/22/192. The contractual start date was in January 2011. The draft report began editorial review in August 2017 and was accepted for publication in March 2018. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).