Fecal immunochemical tests compared with guaiac fecal occult blood tests for population-based colorectal cancer screening

Linda Rabeneck MD MPH FRCPC (Chair)[1], R Bryan Rumble MSc[2], Frank Thompson MD[3], Michael Mills MD CCFP, FCFP[4], Curtis Oleschuk MD FCACB[5], Alexandra Whibley BSc(Hons)[6], Hans Messersmith MPH[7], Nancy Lewis PhD[6]; The FIT Guidelines Expert Panel*

L Rabeneck, RB Rumble, F Thompson, et al; The FIT Guidelines Expert Panel. Fecal immunochemical tests compared with guaiac fecal occult blood tests for population-based colorectal cancer screening. Can J Gastroenterol 2012;26(3):131-147.

Colorectal cancer (CRC) is the second most common cause of cancer deaths in Canadian men and women – accounting for almost 12% of all cancer deaths. In Ontario, it is estimated that 8100 persons were diagnosed with CRC in 2011, and 3250 died from the disease. CRC incidence and mortality rates in Ontario are among the highest in the world. Screening offers the best opportunity to reduce this burden of disease. The present report describes the findings and recommendations of Cancer Care Ontario's Fecal Immunochemical Tests (FIT) Guidelines Expert Panel, which was convened in September 2010 by the Program in Evidence-Based Care. The purpose of the present guideline is to evaluate the existing evidence concerning FIT to inform the decision on how to replace the current guaiac fecal occult blood test with FIT in the Ontario ColonCancerCheck Program. Eleven articles were included in the present guideline, comprising two systematic reviews, five articles reporting on three randomized controlled trials and reports of four other studies. Additionally, one laboratory study was obtained that reported on several parameters of FIT tests that helped to inform the present recommendation. The performance of FIT is superior to the standard guaiac fecal occult blood test in terms of screening participation rates and the detection of CRC and advanced adenoma. Given greater specimen instability with the use of FIT, a pilot study should be undertaken to determine how to implement the FIT in Ontario.

Key Words: *Cancer Care Ontario; Colorectal cancer screening; Fecal immunochemical tests; Guaiac fecal occult blood test*

Les tests immunologiques de sang occulte dans les selles par rapport à la recherche de sang occulte au gaïac pour dépister le cancer colorectal en population

Le cancer colorectal (CCR) est la deuxième cause de décès par cancer en importance chez les hommes et femmes canadiens. Il représente près de 12 % de tous les décès par cancer. En Ontario, on estime que 8 100 personnes ont reçu un diagnostic de CCR en 2011 et que 3 250 en sont morts. L'Ontario affiche l'un des taux d'incidence et de mortalité du CCR les plus élevés au monde. Le dépistage représente la meilleure occasion d'en réduire le fardeau. Le présent rapport décrit les observations et les recommandations des lignes directrices du groupe d'experts d'Action Cancer Ontario sur les tests immunochimiques du sang occulte dans les selles (TIS), qui s'est réuni en septembre 2010 sous l'égide du programme de soins fondés sur des données probantes. Les présentes lignes directrices visent à évaluer les données actuelles au sujet du TIS afin d'étayer la décision sur la manière de remplacer la recherche de sang occulte dans les selles au gaïac par le TIS dans le cadre du Programme ContrôleCancerColorectal de l'Ontario. Onze articles font partie des présentes lignes directrices, y compris deux analyses systématiques, cinq articles sur trois essais aléatoires et contrôlés et des rapports sur quatre autres études. De plus, une étude de laboratoire portait sur plusieurs paramètres du TIS et a contribué à documenter la présente recommandation. Le rendement du TIS est supérieur à celui de la recherche de sang occulte dans les selles au gaïac pour ce qui est des taux de participation au dépistage et de la détection du CCR et de l'adénome avancé. Étant donné la plus grande instabilité des échantillons dans le cadre du TIS, il faudrait entreprendre un projet pilote pour déterminer comment mettre en œuvre le TIS en Ontario.

Colorectal cancer (CRC) is the second most common cause of cancer deaths in Canadian men and women, accounting for almost 12% of all cancer deaths (1). In Ontario in 2011, an estimated 8100 persons were diagnosed with CRC, and 3250 died from the disease. Colorectal cancer incidence and mortality rates in Ontario are among the highest in the world (2). Screening offers the best opportunity to reduce this burden of disease.

The two CRC screening methods recommended by the Canadian Task Force on Preventive Health Care for men and women at average risk for CRC (ie, asymptomatic, 50 years of age or older, and with no other risk factors for CRC) are the fecal occult blood test (FOBT) and flexible sigmoidoscopy (FS) (3). These recommendations are supported by evidence from randomized controlled trials (RCTs). In the 1990s, evidence from RCTs demonstrated that screening with the FOBT (coupled with colonoscopy for those who test positive) is associated with a decrease in CRC mortality and an increase in the proportion of detected cancers that are at Dukes' Stage A (4-6). In 2010, results from the United Kingdom (UK) Flexible Sigmoidoscopy trial also demonstrated that screening with FS is associated with a decrease in CRC mortality (7).

In January 2007, the Ontario Ministry of Health and Long-Term Care announced funding for a province-wide, population-based CRC screening program. The program, 'ColonCancerCheck', uses FOBT for screening individuals at average risk, and colonoscopy as the initial screening test for those at increased risk because of a family history of one or more first-degree relatives diagnosed with CRC. Colonoscopy is also used to investigate screenees with a positive FOBT. Colonoscopy standards were developed by Cancer Care Ontario (CCO)'s Program in Evidence-Based Care (PEBC) to support the ColonCancerCheck Program (8).

[1]*Prevention and Cancer Control, Cancer Care Ontario, Department of Medicine, University of Toronto, Toronto;* [2] *Program in Evidence-based Care, Cancer Care Ontario, Department of Oncology, McMaster University, Hamilton;* [3]*Community Pathology, Cancer Care Ontario, Toronto;* [4]*Hamilton Niagara Haldimand Brant LHIN, Cancer Care Ontario, Hamilton, Ontario;* [5]*Department of Clinical Biochemistry and Genetics, Diagnostic Services of Manitoba, Winnipeg, Manitoba;* [6]*Cancer Care Ontario, Toronto;* [7]*Quality and Methods, Program in Evidence-Based Care, Cancer Care Ontario, Department of Oncology, McMaster University, Hamilton, Ontario. *The complete listing of the FIT Expert Guidelines Panel is presented in Appendix A.*

Correspondence: Dr Linda Rabeneck, Cancer Care Ontario, 620 University Avenue, Toronto, Ontario M5G 2L7. Telephone 416-971-9800 ext 3869, fax 416-971-6888, e-mail linda.rabeneck@cancercare.on.ca

Received for publication December 20, 2011. Accepted December 30, 2011

TABLE 1
Fecal immunochemical tests (FIT) with an active license approved for use by Health Canada*

Manufacturer/distributor	Device	Product description	Numerical or visual result
Eiken†/Polymedco‡	OC-Auto Micro 80 FOB Test System (believed equivalent to OC-Hemodia)	Flat tube, dipstick collection, machine developed	Numerical
Eiken/Polymedco	OC-Sensor DIANA IFOB Test System (Assumed to be the same as OC-Auto Micro)	Flat tube, dipstick collection, machine developed	Numerical
Alfresa Pharma Corporation§/Inverness Medical¶	I-FOBT Hemoglobin NS-Plus	Flat tube, dipstick collection, machine developed	Numerical
Beckman Coulter**	Hemoccult ICT, Immunochemical Fecal Occult Blood Test (also known as Flexsure OBT)	Test card, applicator stick, on-card developed	Visual
Eiken/Polymedco	OC-Light Manual IFOBT	Long cylindrical tube, dipstick collection, test strip developed	Visual
Inverness Medical	Clearview Ultra FOB Test	Long cylindrical tube, dipstick collection, test strip developed	Visual
Medix Biochemica††	Actim Fecal Blood Test	Cylindrical tube, sampling stick that then is put into the tube, development occurs on the stick	Visual
PSS World Medical‡‡	Consult Diagnostic Occult Blood Test Extra Sensitive	Unknown	Insert/instructions not available
Artron§§	One Step Fecal Occult Blood Test	Cylindrical tube, dipstick sampling, developed on a cassette	Visual
IND Diagnostic¶¶/BTNX***	Rapid Response One-Step Fecal Occult Blood Test	Cylindrical tube, dipstick sampling, developed on a cassette	Visual
Tremblay Harrison†††	Minute Lab Fecal Occult Blood Test Device	Cylindrical tube, dipstick sampling, developed on a cassette	Visual
WHPM Bioresearch & Technology‡‡‡	Hemosure Immunological Fecal Occult Blood Test	Cylindrical tube, dipstick sampling, developed on a cassette	Visual
Innovacon§§§	FOB One Step Fecal Occult Blood Test	Cylindrical tube, dipstick sampling, developed on a cassette	Visual

**See Appendix B for complete details from the manufacturers inserts and/or provided literature. †Eiken Chemical Company Ltd, Japan; ‡Polymedco Inc, USA; §Alfresa Pharma Corporation, Japan; ¶Inverness Medical, Canada; **Beckman Coulter Inc, USA; ††Medix Biochemica, Finland; ‡‡PSS World Medical Inc, USA; §§Artron Bioresearch, Canada; ¶¶IND Diagnostic Inc, Canada; ***BNTX Inc, Canada; †††Tremblay Harrison Inc, USA; ‡‡‡WHPM Bioresearch & Technology Inc, USA; §§§Innovacon Inc, USA. FOBT Fecal occult blood test*

Before the launch of ColonCancerCheck, an expert panel was convened by the PEBC to evaluate the evidence concerning existing guaiac FOBT (gFOBT) kits and, based on this evidence, to develop gFOBT Standards for the Ontario CRC Screening Program (9). The standards provided a basis for selecting the gFOBT kit used by the ColonCancerCheck Program and determined the laboratory requirements for the program. The selected gFOBT kits have been in use since April 2008.

At the time the gFOBT Laboratory Standards Expert Panel began its work in the fall of 2006, the fecal immunochemical test (FIT) was undergoing evaluation in various settings. However, the body of evidence was not large, and FIT was not endorsed for CRC screening by a screening guideline from any jurisdiction. However, the expert panel anticipated the need to evaluate the evidence concerning FIT as the body of evidence developed.

The present FIT Guidelines Expert Panel was convened in September 2010 by the PEBC to evaluate the evidence concerning existing FIT kits and, based on this evidence, to set forth FIT guidelines for the ColonCancerCheck Program.

The gFOBT and the FIT are based on different analytical principles. The gFOBT indirectly detects blood in the stool that may be due to bleeding from CRC. The test is based on the oxidation of guaiac (impregnated on the card) by hydrogen peroxide catalyzed by the peroxidase activity of hemoglobin. The disadvantage of this reaction is that it will occur with any peroxidase found in feces (eg, plant peroxidases, heme in red meat) and is affected by certain chemicals (eg, vitamin C) (10). Thus, gFOBTs are not specific for human hemoglobin. gFOBTs may also detect bleeding from any site in the gastrointestinal tract, including the stomach (11). To complete a gFOBT, participants are required to apply six fecal samples (two samples from each of three consecutive spontaneously passed stools) onto test areas (windows) on FOBT cards, a type of sampling referred to in the literature as a 'dry' method. . The gFOBT is visually read by trained laboratory technicians using the naked eye to interpret a visual result.

In contrast, the FIT uses an antibody against human globin – the protein part of hemoglobin. The FIT is specific for human hemoglobin, and is more specific than the gFOBT for bleeding from the distal gut (ie, colon and rectum). To complete a FIT, participants sample one or more stools using various sampling systems, and samples are either applied to a card (dry method) or placed into a vial, a type of sampling referred to in the literature as a 'wet' method. Devices used to collect the stool include wooden sticks and brushes. For some manufacturer's FIT kits, samples are analyzed using automated systems in the laboratory. These systems provide a numerical result and allow for a customized cut-off in hemoglobin concentration to be set to define a positive test. In contrast, other FIT kits are designed as point-of-care devices with prespecified cut-off points used to define a positive result. Similar to gFOBT, they are read by the naked eye, with a positive result indicated by a colour change on a strip. They are designed for doctor's offices or clinics, but can be adapted for use in clinical laboratories for high-volume, population-based screening, albeit using a more manual approach compared with the automated systems.

Although there are many brands of FIT kits available, the focus of the present document is on the 13 FITs that are currently approved for processing in a laboratory setting in Canada (level 2 approval, Health Canada) (Table 1). The expert panel's opinion is that test processing in laboratories, rather than point-of-care processing, is essential for a population-based screening program, which requires quality control protocols in laboratories as well as population-level data collection to monitor program performance.

METHODS

Clinical questions

To inform recommendations regarding how to replace the current gFOBT with a FIT in the population-based CRC screening program in Ontario, the expert panel evaluated existing evidence concerning the following three key aspects of FIT kit use:

TABLE 2
Literature search sources

Source/database	Date searched	Number of hits	Ordered for full–text review	Retained
Systematic search. See Appendix D for strategies used				
MEDLINE	June 17, 2010	227	33	4
EMBASE	June 17, 2010	362	31	0*
Keyword search. Terms used: fecal occult blood test, immunochemical, FIT, FOBT, colorectal cancer, screening				
Canadian organizations				
British Columbia Cancer Agency (www.bccancer.bc.ca)	October 12, 2010	0	–	–
Alberta Cancer Board (www.cancerboard.ab.ca)	October 12, 2010	0	–	–
Saskatchewan Cancer Agency (www.saskcancer.ca)	October 12, 2010	0	–	–
Cancer Care Manitoba (www.cancercare.mb.ca)	October 12, 2010	0	–	–
Cancer Care Nova Scotia (www.cancercare.ns.ca)	October 12, 2010	0	–	–
United States organizations				
NGC (www.guidelines.gov)	June 17, 2010	5	–	–
AHRQ HTA (www.ahrq.gov)	October 12, 2010	1	0	–
ASCO (www.asco.org)	October 12, 2010	0	–	–
NCCN (www.nccn.org)	October 12, 2010	1	0	–
United Kingdom (UK) organizations				
Cochrane Database of Systematic Reviews	June 17, 2010	1	1	1
UK NHS HTA (www.hta.ac.uk)	October 12, 2010	0	–	–
NICE (www.nice.org.uk)	October 12, 2010	2	2†	–
SIGN (www.sign.ac.uk)	October 12, 2010	0	–	–
Cancer UK (www.canceruk.org)	October 12, 2010	0	–	–
Cancer Services Collaborative, Avon, Somerset and Wiltshire (www.aswcs.nhs.uk)	October 12, 2010	0	–	–
NHS (www.nhs.uk)	October 12, 2010	6	0	–
Australian organizations				
National Health & Medical Research Council (www.nhmrc.gov.au)	October 12, 2010	1	0	–
The Cancer Council Australia (www.cancer.org.au)	October 12, 2010	1	0	–
National Cancer Control Initiative (www.canceraustralia.gov.au)	October 12, 2010	0	–	–
State Government of Victoria (www.vic.gov.au)	October 12, 2010	0	–	–
Peter MacCallum Cancer Centre (www.petermac.org)	October 12, 2010	0	–	–
Medical Oncology Group of Australia (www.moga.org.au)	October 12, 2010	0	–	–
New Zealand organizations				
New Zealand Guidelines Group (www.nzgg.org.nz)	October 12, 2010	1	0	–
New Zealand Cancer Control Trust (wwwcancercontrol.org.nz)	October 12, 2010	0	–	–
Obtained through other resources (eg, articles forwarded by panel members, etc)				
Various	Various	NA	NA	6
TOTAL				11

*Data presented as n. *No EMBASE articles remained after MEDLINE duplicates were removed; †Duplicate publications found in MEDLINE search. AHRQ Agency for Healthcare Research & Quality; ASCO American Society of Clinical Oncology; FIT Fecal immunochemical test; FOBT Fecal occult blood test; HTA Health Technology Assessment; NA Not available; NCCN National Comprehensive Cancer Network; NGC National Guidelines Clearinghouse; NHS National Health Service; NICE National Institute for Health and Clinical Excellence; SIGN Scottish Intercollegiate Guidelines Network*

1. FIT performance factors: What are the performance characteristics (sensitivity, specificity, positivity and positive predictive value [PPV]) of FIT when used to detect CRC? (see Appendix C for definitions of the diagnostic parameters).
2. FIT kit usability factors: What FIT kit factors affect acceptability by users (eg, card or 'dry' versus vial, or 'wet' collection FIT, medication use)?
3. Specimen stability: What factors affect specimen stability?

Literature search

The MEDLINE and EMBASE databases were systematically searched for articles assessing FIT screening for CRC published between 1996 and indexed through June 2010. The search strategies used are listed in Appendix D. Additionally, the websites of a large number of agencies and organizations were also searched for evidence, and a listing of all sources searched and the number of articles ordered and retained is presented in Table 2. Expert panel members were also canvassed to ensure that no relevant articles were missed.

In addition to the evidence obtained in the present review, the knowledge obtained from the ColonCancerCheck Program will be considered when making recommendations.

Selection criteria

Eligible sources of information were required to meet the following criteria:

1. Published full reports with information on any of performance, usability or specimen stability factors as listed above.
2. Systematic reviews (SRs), RCTs, other prospective study designs, retrospective study designs and mixed design studies. For the purpose of the present article, SRs, including those that are the evidentiary foundation for clinical practice guidelines or health technology assessments, or similar reports were included provided that they reported in detail (eg, search methods, selection criteria) on a systematic search and summary of the health care literature for articles on a relevant topic.

TABLE 3
Literature search results

Author (reference), year	Fecal tests	Outcomes reported
Randomized controlled trial(s)		
van Rossum et al (19), 2008	Hemoccult II (Beckman Coulter*) versus OC-Sensor (Eiken†)	Primary outcome for which the trial was powered: CRC detection rates. Other outcomes: specificity, positivity, PPV, participation, AA detection rates
van Rossum et al (18), 2009 (*Note: this is further analysis of the same data set used in van Rossum et al [19], 2008)*	Hemoccult II (Beckman Coulter) versus OC-Sensor (Eiken)	Specificity of the OC-Sensor FIT at different cut-off levels in hemoglobin concentration (ng/mL)
Hoffman et al (15), 2010	Hemoccult II (Beckman Coulter) versus OC-Micro (Eiken)	Primary outcome for which the trial was powered: participation (adherence)
Hol et al (17), 2009	Hemoccult II (Beckman Coulter) versus OC-Sensor Micro (Eiken)	Primary outcome for which the trial was powered: participation Other outcomes: specificity, positivity, CRC and AA detection rates, and PPV
Hol et al (16), 2010 (*Note: this is further analysis of the same data set used in 2009 Hol [17], 2009)*	Hemoccult II (Beckman Coulter) versus OC-Sensor Micro (Eiken)	Specificity, positivity and PPV at different cut-off levels in hemoglobin concentration (ng/mL)
Other studies		
Allison et al (20), 2007	Hemoccult SENSA (Beckman Coulter) versus FlexSure OBT/Hemoccult ICT (Beckman Coulter)	Primary outcome: advanced neoplasia (defined as CRC or AA) in the distal colon. Other outcomes: specificity, positivity and PPV in testing 3 consecutive bowel movements
Grazzini et al (22), 2009	OC-Hemodia and OC-Sensor Micro (Eiken)	Positivity and PPV comparing a 1 versus 2 day sampling strategy at different cut-off levels in hemoglobin concentration (ng/mL)
Park et al (23), 2010	Hemoccult II (Beckman Coulter) versus OC-Sensa Micro (Eiken)	Primary outcome: detection of advanced neoplasia. Other outcomes: specificity, positivity and sensitivity in testing 3 consecutive bowel movements at different cut-off levels in hemoglobin concentration (ng/mL)
Grazzini et al (21), 2010	OC-Sensor (Eiken)	Effect of seasonal temperature variation on positivity and PPV

**Beckman Coulter Inc, USA; †Eiken Chemical Company, Ltd, Japan. AA Advanced adenoma; CRC Colorectal cancer; FIT Fecal immunochemical test; FOBT Fecal occult blood test; PPV Positive predictive value*

3. Reports published in English.
4. Reports that evaluated at least one FIT kit that is licensed by Health Canada for use in Canada.
5. Studies that did not include symptomatic participants.

Quality assessment of included evidence

An assessment of study quality was performed for all the included evidence. For RCTs, no specific instrument was used, but items such as randomization, sample size estimates and power calculation, and funding sources were reported on. The expert panel recognized that, due to the nature of the studies being examined, blinding to the intervention was not always possible and, therefore, the lack of blinding was not considered to be a methodological flaw, nor was lack of a reported period of follow-up.

For the other evidence types, the Quality Assessment of Studies of Diagnostic Accuracy included in Systematic Reviews (QUADAS) tool was used where appropriate (12). The QUADAS tool is a 14-item questionnaire intended to assess primary studies of diagnostic utility for SRs. The QUADAS instrument can only be used to assess studies of diagnostic utility in which one test is compared with another (typically, the gold standard). For diagnostic studies without a comparator, no formal quality assessment was planned.

RESULTS

A total of 11 articles were retained, comprising two SRs (13,14), five articles reporting on three RCTs (15-19) and articles regarding four other studies (20-23). The two SRs retrieved in the literature search, Whitlock et al (14) and Mujoomdar et al (13), identified studies that were also retrieved in the literature search. Whitlock et al (14) was a SR commissioned by the United States (US) Preventive Services Task Force (USPSTF) to provide updated recommendations on CRC screening with the development of newer tests; the relevant data from this article were incorporated into the present review and referenced from the primary sources. Mujoomdar et al (13) was a SR commissioned by the Canadian Agency for Drugs and Technologies in Health (CADTH) to analyze the available evidence on the accuracy of and compliance to FIT compared with gFOBT in CRC screening. Again, the relevant data from this article have been incorporated into the present review from the primary sources.

A laboratory study by Lamph et al (24), conducted on behalf of the National Health Service in the UK, reported on an independent assessment of several parameters of FIT tests, especially temperature stability. Of the three tests evaluated, one, the OC-Sensor (Eiken Chemical Company, Ltd, Japan) product, is approved in Canada. All of the findings are reported in the Results section.

Four additional articles were also obtained and retained for discussion purposes (25-28). Two were US studies that did not meet the inclusion criteria: Rex et al (28) for the American College of Gastroenterology, Levin et al (27) for the American Cancer Society, US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. The article by Halloran et al (25), for the European Guidelines for Quality Assurance in Colorectal Cancer Screening, did meet the inclusion criteria, but the full publication appeared after the end date of our search (June 2010). The fourth article, from the Health Council of the Netherlands (26), did not meet our inclusion criteria. Although retained for discussion purposes, no quality assessment was performed on these reports because no formal adaptation was planned.

Quality of included studies

As described previously, RCTs were assessed for quality according to the following criteria: randomization, details of the statistical analysis, expected effect size and details of the statistical power calculation, differences in patient characteristics and funding sources (Table 3).

The RCT reported by Hoffman et al (15) randomly assigned patients to either FIT or gFOBT using an online random number generator. The primary outcome was screening adherence (subsequently referred to herein as 'participation'). The two groups were compared using the *t* test or the Wilcoxon two-sample test (for continuous variables) and the χ^2 test (for categorical variables). Participation (defined as test completion within 90 days) was compared using χ^2 and multivariate logistic regression (adjusted stepwise for demographics [age,

sex, race/ethnicity and clinic site], previous testing and comorbidities). The expected effect size was a 10% difference in screening participation. Statistical power was determined using the results of a pilot study showing 40% participation with gFOBT; therefore, to detect a 10% difference with 80% power, a minimum of 800 participants was required. However, the actual number of participants was 404, and it was unclear from the report why the minimum number of participants was not included. Differences in patient characteristics were reported, with no differences between groups detected. The Department of Veterans Affairs (US) was the source of funding.

The RCT reported by Hol et al in two publications (16,17) randomly assigned participants to three groups, FIT, gFOBT and FS, on a 1:1:1 basis using a computer-generated algorithm. Participants were stratified according to age, sex and socioeconomic status (SES). The primary outcome was participation. Detection of advanced neoplasia (AN) (defined as those with either CRC or an advanced adenoma [AA]) was the secondary outcome. The three groups were compared using the χ^2 test to detect differences in proportions, the *t* test to detect differences in means between screening strategies, and univariate logistic regression to detect differences in participation rates among the three screening strategies, with multivariate modelling used to investigate possible interactions. The expected effect size was a 2% difference in participation rates among the three screening strategies and a 2.5% difference in participation rates between a maximum of three equal-sized subgroups per arm. Based on an expected participation rate of 50%, the sample size was calculated using 80% power (the exact sample size needed was not explicitly stated). Differences in patient characteristics were reported, and were detected for the following comparisons: sex (more women in FIT and FS arm), age 50 to 59 years (more in FIT and FS arm), age 65 to 74 years (more in FS arm), SES middle (more in gFOBT arm), SES high (more in FIT and FS arm), urban (more in gFOBT arm) and rural (more in gFOBT arm compared with FIT and FS; more in FIT compared with FS) (all differences reported $P<0.05$). The Dutch Cancer Society, Dutch Ministry of Health (Health Care Prevention Program), Olympus Medical Systems Europe GmbH (Germany) and Eiken Chemical Company, Ltd (Japan) were sources of funding.

The RCT reported by van Rossum et al in two publications (18,19) randomly assigned participants according to postal address to either FIT (OC-Sensor [Eiken Chemical Company, Ltd, Japan] [n=10,322]) or gFOBT (Hemoccult II [Beckman Coulter Inc, USA] [n=10,301]) using a study-specific randomization program. The primary outcome was CRC detection. Differences for participation, positivity, detection, PPV and specificity were calculated using a two-group χ^2 test reported with a 95% CI. The expected effect size was a 0.3% difference in CRC detection. To detect this difference at 80% power, a sample size of 10,000 in each group was required, but was unclear whether this number refers to persons invited or to persons that participated. Although 20,623 participants were randomized and invited, 10,993 participated (4836 gFOBT; 6157 FIT). Differences in patient characteristics were reported, with no differences observed. The Netherlands Organization for Health Research & Development was the funding source.

The expert panel did not detect major methodological flaws in the three RCTs included in the present review.

For the studies on diagnostic accuracy, the QUADAS instrument was used for quality assessment (12). QUADAS requires a test of diagnostic accuracy with an appropriate comparator, and neither of the Grazzini et al (21,22) studies compared FIT with either colonoscopy or a gFOBT. Only the studies reported by Allison et al (20) and Park et al (23) were assessed for quality using QUADAS. A summary of the QUADAS results follows.

Both Allison et al (20) and Park et al (23) studied a group that was representative of the types of participants that would receive screening in practice and had clearly defined selection criteria. The reference standards used for each study were different, with Allison et al comparing FIT with gFOBT, and with Park et al comparing FIT with colonoscopy. In both studies, the reference standard and the index test were performed in a time period sufficiently short to ensure the target condition would not have changed. In Allison et al, colonoscopy was not used as the comparator to FIT, although participants with a positive test were referred for colonoscopy, and those with a negative test were referred for FS. In Park et al, FIT results were compared directly with colonoscopy results. For both studies, the index test was independent of the reference standard used. The methods describing the use of the reference standard (gFOBT) used in the study by Allison et al were not described well enough in the methods section to enable replication of the procedure by others, while the Park et al study fully described the colonoscopy procedure used, thereby enabling replication. Neither study reported whether the results of the index test or the reference standard were assessed independently. The clinical data collected in the study reported by Allison et al would be available in clinical practice, but this was not clear from the Park et al study because details of the clinical data collected were not provided. Both studies reported on uninterpretable results and how withdrawals from the study were handled and reported. Because both Park et al and Allison et al were studies that reported on the diagnostic utility of FIT versus other diagnostic tests, no power calculation was described, and no primary outcome was identified. Details of the QUADAS assessment appear in (Appendix E).

RESULTS: EVIDENCE CONCERNING THE THREE KEY ASPECTS OF FIT KIT USE

FIT performance factors

Literature search results: ***Comparing the performance of FIT with gFOBT:*** Here we report on the four articles that provided a comparison of FIT performance (at the manufacturer's recommended cut-off level of 100 ng/mL of hemoglobin) to gFOBT (17,19,20,23). Although the focus of the expert panel was on the performance of FIT for detecting CRC, when included articles also reported on the detection of AA (precancerous lesions that have the potential to develop into CRC if left untreated), the data on AA were also included.

The RCT conducted by van Rossum et al (19) was a study of 20,623 men and women 50 to 75 years of age. The study compared the performance of Hemoccult II over three days in 10,301 participants with the performance of one OC-Sensor sample in 10,322 participants. If one of the samples tested positive either through Hemoccult II giving a visual colour reaction or the OC-Sensor output yielding a numerical value >100 ng/mL as recommended by the manufacturer, participants were referred for follow-up colonoscopy. Participants with negative tests did not undergo follow-up colonoscopy. Tests were returned to the laboratory via the postal system and, if not returned immediately, participants were advised to refrigerate the sample. Once at the laboratory, tests were stored at 4°C if not developed immediately. Of the returned tests, 75% were developed within two days of receipt at the laboratory, and 99.6% of tests were developed within six days. The definition of AA used by van Rossum et al (19) was adenomas ≥10 mm with high-grade dysplasia or with a villous component ≥20%.

The RCT conducted by Hol et al (17) was a study of 10,011 men and women 50 to 74 years of age. The results for 5004 people who completed the Hemoccult II gFOBT over three days were compared with the results for 5007 people who completed one sample for the OC-Sensor FIT. Tests were returned to the laboratory via the postal system. Positivity was defined as at least one positive panel identified by a visual colour reaction from Hemoccult II or a numerical output >100 ng/mL for OC-Sensor. All persons with a positive Hemoccult II test were referred for follow-up colonoscopy, as were those who received a numerical output >50 ng/mL with the OC-Sensor test. No information was provided about the time between sampling and test development. The definition of AA used by Hol et al (17) was adenomas ≥10 mm with high-grade dysplasia or with a villous component ≥25%.

The study conducted by Park et al (23) enrolled a total of 770 men and women 50 to 75 years of age who underwent screening colonoscopy

TABLE 4
Performance characteristics of fecal immunochemical test (FIT) compared with guaiac fecal occult blood test (gFOBT)

Author (ref), year	Study population	Comparisons	Sensitivity, %	Specificity, %	Positivity, %	PPV, %
van Rossum et al (19), 2008	20,623 participants 50–75 years of age	OC-Sensor* FIT (n=10,322) vs Hemoccult II† gFOBT (n=10,301)	Not reported	CRC FIT: 95.8, gFOBT: 98.1; AA[1] FIT: 97.1, gFOBT: 98.7	FIT: 5.5, gFOBT: 2.4	CRC FIT: 8.6, gFOBT: 10.7; AA[1] FIT: 37.9 gFOBT: 39.8
Hol et al (17), 2009	10,011 participants 50–74 years of age	OC-Sensor FIT (n=5007) vs Hemoccult II gFOBT (n=5,004)	Not reported	CRC FIT: 95.8, gFOBT: 97.6; AA[2] FIT: 97.8, gFOBT: 98.5	FIT: 4.8, gFOBT:2.8	CRC FIT: 10, gFOBT:10; AA[2] FIT: 53, gFOBT: 45
Park et al (23), 2010	770 participants 50–75 years of age completed both tests concurrently	OC-Sensor Micro* FIT vs Hemoccult II gFOBT	CRC FIT: 92.3 gFOBT: 30.8; AA: FIT: 33.9 gFOBT: 13.6	CRC FIT: 90.1, gFOBT: 92.4; AA[4] FIT: 90.6, gFOBT: 92.4	FIT: 11.2, gFOBT: 7.9	CRC FIT: 12.8, gFOBT: 6.7; AA[4] FIT: 23.3, gFOBT: 13.1
Allison et al (20), 2007	5932 participants ≥50 years of age completed both tests concurrently	FlexSure OBT†/ Hemoccult ICT† FIT vs Hemoccult SENSA† gFOBT	CRC FIT: 81.8, gFOBT: 64.3; AA: FIT: 29.5, FOBT: 41.3	CRC FIT: 96.9, gFOBT: 90.1; AA[3] FIT: 97.3, gFOBT: 90.6	FIT: 3.2, gFOBT: 10.1	CRC FIT: 5.2, gFOBT: 1.5; AA[3] FIT: 19.1, gFOBT: 8.9

**Eiken Chemical Company, Ltd, Japan; †Beckman Coulter Inc, USA. AA Advanced adenoma (AA[1] Adenoma ≥10 mm with high-grade dysplasia or with a villous component ≥20%; AA[2] Adenoma ≥10 mm with high-grade dysplasia or with a villous component ≥25%; AA[3] Tubular, villous, or tubulovillous adenomas ≥10 mm; AA[4] Tubular adenomas ≥10 mm or tubulovillous or villous adenomas, or those with high-grade dysplasia regardless of size); CRC Colorectal cancer; NR Not reported; ref Reference; vs Versus. All results at manufacturers' suggested cut-off (100 ng/mL).*

in the week after completing the fecal testing. The primary outcome was AN (defined as either CRC or AA) in the colon or rectum (if the cecum was not reached, the patient was excluded from analysis). The study compared three days of sampling using the Hemoccult II gFOBT with three days of sampling using the OC-Sensor FIT. A positive test was considered to be either one of the Hemoccult II samples exhibiting a visual colour reaction or a numerical output from OC-Sensor >100 ng/mL in at least one sample. No information was provided regarding how samples were returned to the laboratory. All Hemoccult II samples were developed on the day of receipt in the laboratory, and OC-Sensor samples were stored at 4°C until sent to the central analysis centre within two days and processed immediately. All samples were developed within two weeks of the first sample collection date. All participants, regardless of result, underwent colonoscopy, allowing for the measurement of sensitivity. The definition of AA used by Park et al (23) was tubular adenomas ≥10 mm or tubulovillous or villous adenomas, or those with high-grade dysplasia regardless of size.

In the study conducted by Allison et al (20), 7394 men and women 50 years of age or older completed both a gFOBT (Hemoccult SENSA [Beckman-Coulter, USA]) and a FIT (Hemoccult ICT [Beckman-Coulter, USA]). Hemoccult SENSA differs from the standard gFOBT (eg, Hemoccult II) because it is more sensitive for detecting CRC (29,30). The primary outcome was AN (defined as either CRC or AA) of the distal or left colon (rectum, sigmoid, descending). Each participant collected a sample each day for three days, and the samples were tested using both the gFOBT and the FIT. Originally, the Hemoccult ICT was developed only if the Hemoccult SENSA had tested positive in at least one of three samples taken. This was changed during the study so that the three Hemoccult ICT samples were developed regardless of the Hemoccult SENSA result. Tests were returned to the laboratory via the postal system and were developed within five days of the first sample. All Hemoccult SENSA tests were developed at least three days after the first sampling date. All Hemoccult ICT were developed within 14 days of the first sampling date. All participants whose stools tested positive were referred for follow-up colonoscopy, and all participants whose stools tested negative were referred for FS, enabling the investigators to compare the sensitivity of gFOBT and FIT for the detection of CRC in the distal colon. The definition of AA used by Allison et al was tubular, villous, or tubulovillous adenomas ≥10 mm.

Table 4 summarizes data extracted from these studies with respect to the performance characteristics of fecal testing, including sensitivity, specificity, positivity and PPV.

In the RCT conducted by Van Rossum et al (19), positivity was statistically significantly higher for FIT than for gFOBT (FIT 5.5% versus gFOBT 2.4% [P<0.01]). The specificity for the detection of CRC and AA was significantly lower for FIT than for gFOBT (CRC FIT 95.8% versus 98.1% [P<0.01]; AA FIT 97.1% versus gFOBT 98.7% [P<0.01]). The difference in PPV for both CRC and AA was not statistically significant when comparing FIT (CRC 8.6%; AA 37.9%) with gFOBT (CRC 10.7%; AA 39.8%). Although the authors were unable to assess sensitivity, they reported on the percentage of persons in whom CRC and AA were detected in each arm of the trial. Using an intention-to-screen analysis, 0.6% of those in the gFOBT arm had either a CRC or AA detected, compared with 1.4% in the FIT arm. This difference was statistically significant, as was the difference in the per protocol analysis.

In the RCT conducted by Hol et al (17), positivity was statistically significantly higher for FIT than for gFOBT (FIT 4.8% versus gFOBT 2.8% [P<0.05]). The specificity for CRC when using FIT (95.8%) was slightly lower than that for gFOBT (97.6%); however, this difference was not statistically significant. The specificity of FIT for AA was statistically significantly lower than that of gFOBT (FIT 97.8% versus gFOBT 98.5%; P<0.05). The difference in PPV for both CRC and AA was not statistically significant when comparing FIT (CRC 10%; AA 53%) with gFOBT (CRC 10%; AA 45%). Although the authors were unable to assess sensitivity, they reported on the percentage of persons in whom CRC and AN were detected in each arm of the trial. Of those in the gFOBT arm, 0.3% had a CRC detected and 1.2% had AN detected, compared with the FIT arm in which 0.5% had CRC detected and 2.5% had AN detected. The difference was statistically significant for the detection of AN but not for CRC.

In the study conducted by Park et al (23), positivity was slightly higher for FIT (11.2%) than for gFOBT (7.9%), but this difference was not statistically significant. The difference in specificity for both CRC and AA was not statistically significant when comparing FIT (CRC 90.1%; AA 90.6%) with gFOBT (CRC 92.4%; AA 92.4%). Again, the difference in PPV for both CRC and AA was not statistically significant when comparing FIT (CRC 12.8%; AA 23.3%) with gFOBT (CRC 6.7%; AA 13.1%). The sensitivity for detecting CRC

TABLE 5
Test characteristics for the detection of colorectal cancer (CRC) using fecal immunochemical tests in multiple sampling

Author (reference), year	Study population	Test type	Comparison	Positivity, %	PPV for CRC, %
Grazzini G et al (22), 2009	20,596 participants 50–69 years of age	OC-Hemodia*	1 test positive	4.5	6.9
			At least 1 test positive	6.7	5.7
			2 tests positive	2.3	10.4

**Eiken Chemical Company, Ltd, Japan. PPV Positive predictive value*

was statistically significantly increased when using FIT compared with gFOBT (FIT 92.3% versus gFOBT 30.8% [P<0.01]). The sensitivity of FIT (33.9%) compared with gFOBT (13.6%) for detecting AA was significantly higher (P<0.05).

In the study conducted by Allison et al (20), positivity was statistically significantly lower for FIT than the sensitive gFOBT used in the study (FIT 3.2% versus gFOBT 10.1% [P<0.01]). The specificity for both CRC and AA was statistically significantly higher for FIT than for gFOBT (CRC FIT 96.9% versus gFOBT 90.1% [P<0.01]; AA FIT 97.3% versus gFOBT 90.6% [P<0.01]). The PPV for both CRC and AA was also statistically significantly higher for FIT compared with gFOBT (CRC FIT 5.2% versus gFOBT 1.5% [P<0.01]; AA FIT 19.1% versus gFOBT 8.9% [P<0.01]). The difference in sensitivity for detecting CRC and AA was not statistically significant for FIT (CRC 81.8%; AA 29.5%) compared with gFOBT (CRC 64.3%; AA 41.3%).

In summary, the sensitivity of FIT for detecting CRC and AA compared with a standard gFOBT, which was assessed in only one study, is superior. In the two Dutch RCTs (17,19), specificity was decreased for CRC and AA when using FIT compared with gFOBT. On the other hand, these two studies reported higher AN detection rates for FIT compared with gFOBT. The PPV for detecting CRC and AA using FIT is not different from the standard gFOBT. In general, the positivity rates for FIT using the manufacturer's standard cut-off level in hemoglobin concentration are higher than for Hemoccult II.

Single-sample testing compared with multiple-sample testing using FIT: Only one study using FIT compared the results of taking multiple samples from consecutive stools with taking one sample from one stool (22). The data from this study are summarized in Table 5.

In this study (Grazzini et al [22]), a single daily sample was compared with testing two samples taken from consecutive bowel movements, using OC-Hemodia (Eiken Chemical Company, Ltd, Japan) at the manufacturer's recommended cut-off level in the hemoglobin concentration of 100 ng/mL. While the samples were taken from consecutive bowel movements, the results were reported as a comparison of a one-day sampling versus a two-day sampling strategy, and both samples were required to be positive to be considered a positive result. Positivity was statistically significantly higher when using a one-day strategy compared with a two-day strategy (4.5% versus 2.3% [P<0.01]). When the definition of a positive result was changed in the two-day strategy to at least one positive test giving a positive result, overall positivity was statistically significantly higher with the two-day strategy (two-day, at least one sample positive: 6.7% versus one-day strategy: 4.5% versus two-day, both samples positive: 2.3% [P<0.01]). There was no statistically significant difference in the PPV with either of the sampling strategies (one-day 6.9%; two-day 10.4%) or when at least one positive test in a two-day strategy was considered positive overall (5.7%). Specificity and sensitivity were not reported.

In summary, positivity rates were affected by the sampling strategies used. As expected, a two-day strategy in which both tests are required to be positive resulted in the lowest positivity rate, while a two-day strategy in which only one test was required to be positive resulted in the highest positivity rate. A one-day strategy resulted in an intermediate positivity rate.

Performance of FIT at different cut-off levels: Here we report on the performance of FIT at multiple hemoglobin concentration cut-off levels that differ from the manufacturer's recommendations. Results from four articles comprising two RCTs (17,18) and two other studies (22,23) are summarized in Table 6.

The RCT reported by van Rossum et al (18) recorded an increasing trend for the specificity of detecting CRC and AA as a combined outcome as the hemoglobin concentration cut-off level increased, but no statistical test results were reported.

In the RCT by Hol et al (17), there was a statistically significant increase in specificity and PPV for detecting both CRC and AA as the hemoglobin concentration cut-off level increased. Positivity was statistically significantly decreased with an increase in hemoglobin concentration cut-off level. FIT was superior to FOBT for CRC detection at 50 ng/mL and 75 ng/mL and for AA detection at 50 ng/mL, 75 ng/mL and 100 ng/mL (P<0.05 for all).

The study by Grazzini et al (22) reported that positivity decreased as the cut-off level increased, while PPV increased. The statistical significance of the differences was not reported. The definition of AA used by Grazzini et al was any adenoma ≥10 mm, and/or a villous component ≥21%, and/or severe dysplasia.

The study by Park et al (23) reported on the effect of increasing the hemoglobin concentration cut-off level on FIT sensitivity. The authors reported that from ≥50 ng/mL to ≥100 ng/mL sensitivity for CRC is unchanged at 92.3% but that above a 100 ng/mL cut-off level sensitivity decreases to 84.6%. For AA, there is a decreasing trend for sensitivity as the hemoglobin concentration cut-off level increases. This study also reports an increasing specificity for detecting CRC and AA as the hemoglobin concentration cut-off level increases, but does not report whether these differences are significant. The study data provided for positivity and PPV were insufficient to assess the effect of increasing the hemoglobin concentration cut-off levels.

In summary, these four studies showed that increasing the hemoglobin concentration cut-off level decreased the positivity rate and increased specificity and PPV. In addition, one study reported that increasing the cut-off level above 100 ng/ml decreased sensitivity.

Information provided in test kit instructions results: No further information on the outcomes of interest was identified in the manufacturer inserts and/or documentation.

FIT kit usability

Literature search results: ***How does FIT compare with gFOBT in user acceptability?*:** Three RCTs (15,17,19) that reported comparative data for screening participation rates using FIT versus gFOBT are summarized in Table 7.

In the RCT conducted by van Rossum et al (19), one-half of the study population was given the gFOBT Hemoccult II (n=10,301) and the other one-half was given the FIT OC-Sensor (n=10,322) to complete. The Hemoccult II test required two samples from a stool on three separate days and involved the smearing of feces onto a card using an applicator stick that then had to be discarded. The OC-Sensor test required one sample from one day and involved scraping the stool sample with a probe that was then inserted into a vial of buffer solution. No dietary or medication restrictions were imposed during the study.

In the RCT conducted by Hoffman et al (15), one-half of the study population was given the gFOBT Hemoccult II (n=202) to complete, and the other one-half was given the FIT OC-Auto (n=202) to complete. The Hemoccult II test required one sample on three separate days and involved the smearing of feces onto a card using an applicator stick that then had to be discarded. The OC-Auto test required two samples

TABLE 6
Performance characteristics of fecal immunochemical tests at different cut-offs in hemoglobin concentration

Author (ref), year	Study population	Cut-off value	Sensitivity		Specificity		Positivity	PPV	
van Rossum et al (18), 2009	428 participants 50–75 years of age using OC-Sensor*				CRC + AA[1]			CRC	AA[1]
		≥50 ng/mL	NR		96.0		8.5	NR	NR
		≥75 ng/mL	NR		97.1		NR	NR	NR
		≥100 ng/mL	NR		97.8		NR	NR	NR
		≥125 ng/mL	NR		98.1		NR	NR	NR
		≥150 ng/mL	NR		98.3		NR	NR	NR
		≥175 ng/mL	NR		98.4		NR	NR	NR
		≥200 ng/mL	NR		98.6		NR	NR	NR
		≥225 ng/mL	NR		98.7		NR	NR	NR
Hol et al (17), 2009	5007 participants 50–74 years of age using OC-Sensor				CRC	AA[2]		CRC	AA[2]
		≥50 ng/mL	NR		92.9	95.5	8.1	7	42
		≥75 ng/mL	NR		95	97.2	5.7	9	49
		≥100 ng/mL	NR		95.8	97.8	4.8	10	53
		≥125 ng/mL	NR		96.3	98.2	4.1	11	57
		≥150 ng/mL	NR		96.6	98.4	4	11	60
		≥175 ng/mL	NR		97	98.7	3.6	12	63
		≥200 ng/mL	NR		97.1	98.8	3.5	12	62
Grazzini et al (22), 2009	20,596 participants 50–69 years of age using OC-Hemodia* and OC-Sensor				CRC	AA[3]		CRC	AA[3]
		≥80 ng/mL	NR		NR	NR	5.5	5.9	NR
		≥100 ng/mL	NR		NR	NR	4.5	6.9	NR
		≥120 ng/mL	NR		NR	NR	4.0	7.6	NR
Park et al (23), 2010	770 participants 50–75 years of age using OC-Sensor Micro*		CRC	AA[4]	CRC	AA[4]	–	CRC	AA[4]
		≥50 ng/mL	92.3	44.1	87.2	88.3	NR	NR	NR
		≥75 ng/mL	92.3	37.3	89.0	89.7	12.2	NR	NR
		≥100 ng/mL	92.3	33.9	90.1	90.6	11.2	12.8	23.3
		≥125 ng/mL	84.6	28.8	91.3	91.6	NR	NR	NR
		≥150 ng/mL	84.6	27.1	91.9	92.1	NR	NR	NR

*Data presented as % unless otherwise indicated. *Eiken Chemical Company Ltd, Japan; AA Advanced adenoma (AA[1] Adenomas ≥10 mm with high-grade dysplasia or with a villous component ≥20%; AA[2] Adenomas ≥10 mm with high-grade dysplasia or with a villous component ≥25%; AA[3] Adenoma ≥10 mm and/or a villous component ≥21% and/or severe dysplasia; AA[4] tubular adenomas ≥10 mm or tubulovillous or villous adenomas, or those with high-grade dysplasia regardless of size); CRC Colorectal cancer; NR Not reported; PPV Positive predictive value; ref Reference*

TABLE 7
Screening participation rates

Author (reference), year	Study population	Comparisons	Participation rate, %
Van Rossum et al (19), 2008	20,623 participants 50–75 years of age	Hemoccult II (n=10,301) vs OC-Sensor* (n=10,322)	FIT: 59.6, gFOBT: 46.9; P<0.01
Hoffman et al (15), 2010	404 participants, two samples taken	Hemoccult II (n=202) vs OC-Auto* (n=202)	FIT: 68, gFOBT: 55; P=0.01
Hol et al (17), 2009	10,011 participants 50–74 years of age	Hemoccult II (n=5004) vs OC-Sensor* (n=5007)	FIT: 61.5, gFOBT: 49.5; P<0.05

**Eiken Chemical Company Ltd, Japan. FIT Fecal immunochemical test; gFOBT Guaiac fecal occult blood test; vs Versus*

from two consecutive stools and involved scraping the stool sample with a probe that was then inserted into a vial of buffer solution. The Hemoccult II study population were instructed to avoid nonsteroidal anti-inflammatory drugs, as well as rare meat, foods containing peroxidase and vitamin C during the three days of sampling.

In the RCT conducted by Hol et al (17), one-half the study population was given the gFOBT Hemoccult II (n=5004), and the other one-half was given the FIT OC-Sensor (n=5007) to complete. The Hemoccult II test required one sample on three separate days and involved the smearing of feces onto a card using an applicator stick that then had to be discarded. The OC-Sensor test required one sample from one day and involved scraping the stool sample with a probe that was then inserted into a vial of buffer solution. No dietary or medication restrictions were imposed during the study

In summary, all three RCTs reported significantly higher participation rates with FIT compared with gFOBT. This increased participation rate for FIT may be attributed to a simpler collection method with fewer samples required, less stool handling and no need for stick disposal. In addition, Hoffman et al (15) required dietary and medication restrictions in the gFOBT group, which could have led to decreased participation.

Information provided in test kit instructions results

Number and timing of samples collected: The manufacturers of most of the approved tests recommend that one sample be collected from one bowel movement. The instructions for the Hemoglobin NS-Plus test from Alfresa recommend two samples collected across two days, and those for the Hemoccult ICT test from Beckman Coulter recommend three samples across three days. These results are summarized in Table 8.

Diet and medication restrictions: Three of the 13 FIT kits provided instructions advising restrictions on alcohol and discontinuation of acetylsalicylic acid and similar medications for 48 h before stool sampling. These results are summarized in Table 8.

Specimen stability

Literature search results: The stability of hemoglobin in the fecal sample is an issue that has arisen with the vial collection method that characterizes the majority of FITs. Temperature and time are the two variables that play a role in the stability of the stool specimen, requiring consideration when implementing a population-based CRC screening program using FIT. Two articles reported on the stability of

TABLE 8
Diet and medication restrictions

Manufacturer/ distributor	Device	Diet/medication restrictions	Number of samples
Eiken*/Polymedco†	OC-Auto Micro 80 FOB Test System (believed equivalent to OC-Hemodia)	None noted	1/day
Eiken/Polymedco	OC-Sensor DIANA IFOB Test System (Assumed the same as OC-Micro)	None noted	1/day
Alfresa Pharma Corporation‡/Inverness Medical§	I-FOBT Hemoglobin NS-Plus	None noted	1/day for 2 days
Beckman Coulter¶	Hemoccult ICT, Immunochemical Fecal Occult Blood Test (also known as Flexsure OBT)	None noted	1/day for 3 days
Eiken/Polymedco	OC-Light Manual IFOBT	Insert/instructions not available	1/day
Inverness Medical	Clearview Ultra FOB Test	None noted	1/day
Medix Biochemica**	Actim Fecal Blood Test	Insert/instructions not available	Unclear but appears to be 1/day
PSS World Medical††	Consult Diagnostic Occult Blood Test Extra Sensitive	Insert/instructions not available	Unknown
Artron‡‡	One Step Fecal Occult Blood Test	None noted	1/day
IND Diagnostic§§/ BTNX¶¶	Rapid Response One-Step Fecal Occult Blood Test	Alcohol, acetylsalicylic acid, and similar medications should be discontinued for 48 h before sample collection	1/day
Tremblay Harrison***	Minute Lab Fecal Occult Blood Test Device	Alcohol, acetylsalicylic acid, and similar medications should be discontinued for 48 h before sample collection	1/day
WHPM Bioresearch & Technology†††	Hemosure Immunological Fecal Occult Blood Test	None noted	1/day
Innovacon‡‡‡	FOB One Step Fecal Occult Blood Test	Alcohol, acetylsalicylic acid, and similar medications should be discontinued for 48 h before sample collection	1/day

**Eiken Chemical Company, Ltd, Japan; †Polymedco Inc, USA; ‡*Alfresa Pharma *Corporation, Japan; §Inverness Medical, Canada; ¶Beckman Coulter Inc, USA; **Medix Biochemica, Finland; ††PSS World Medical Inc, USA; ‡‡Artron Bioresearch, Canada; §§IND Diagnostic Inc, Canada; ¶¶BNTX Inc, Canada; ***Tremblay Harrison Inc, USA; †††WHPM Bioresearch & Technology Inc, USA; ‡‡‡Innovacon Inc, USA. FOB Fecal occult blood*

the sample in varying temperatures (21,24); Lamph et al (24) also examined the effect of time at selected temperatures.

The study reported by Grazzini et al (21) indirectly measured the effects of ambient temperature and moisture on collected samples in a screening study in an Italian population across different seasons. In this study, the PPV for the detection of CRC and AA did not vary significantly from season to season, ranging from 24% to 26%. However, in a logistic regression analysis that adjusted for age, sex and history of screening (first or repeated test), the odds of having a positive screening test was significantly lower in summer (OR 0.83 [95% CI 0.76 to 0.90]), autumn (OR 0.88 [95% CI 0.83 to 0.94]) and spring (OR 0.90 [95% CI 0.85 to 0.96]) compared with the probability in winter. When the analysis used average ambient temperature in the five to 11 days before the test analysis, an increase of 1°C resulted in 0.7% reduced odds of a positive FIT (OR 0.993 [95% CI 0.989 to 0.996]). The authors concluded that in summer the probability of detecting CRC or AA is approximately 13% lower than in winter. This study reported a mean of 11 days between sample collection and laboratory development but did not analyze the effects of time and temperature together.

Lamph et al (24) conducted an independent evaluation of the temperature stability of the OC-Sensor product and subsequently verified the manufacturer's reported temperature stability values. Table 9 summarizes these data.

Information provided in test kit instructions results

Table 10 provides details on temperature stability and storage times and conditions for the 13 Health Canada-approved FIT kits. According to the information provided by the manufacturer with three of the FITs, specimens are stable for seven days (I-FOBT at 25°C, Hemoccult ICT at 15°C to 30°C and Clearview UltraFOB at 2°C to 8°C) and with two of the FITs samples are stable for ≥15 days (OC-Auto at 15°C to 30°C, OC-Light at 15°C to 30°C).

TABLE 9
Temperature stability* for the OC-Sensor† fecal immunochemical test kit

Storage temperature, °C	Manufacturer claimed stability versus measured stability, days	
−18 to −24	Claimed	10–14
	Measured	Agree
4 to 8	Claimed	7
	Measured	Agree
23 to 26	Claimed	3
	Measured	Agree
29 to 34	Claimed	No claim made
	Measured	<3

**As measured by Lamph et al (24); †Eiken Chemical Company, Ltd, Japan.*

Implementing FIT in population-based CRC screening programs: *Recommendations from other jurisdictions:* Two guidelines from the US were identified. A guideline by Rex et al (28) for the American College of Gastroenterology recommended annual FIT over card-based gFOBT because FIT has both superior test characteristics and adherence rates for the detection of CRC. A guideline by Levin et al (27) for the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology stated that annual testing with either a high-sensitivity gFOBT or FIT in both male and female participants aged 50 years and older are both acceptable options for CRC screening.

In the UK, the National Health Service Evaluation Report on "Immunochemical faecal occult blood tests" (24) provided a comparative analysis of three FIT devices available in the UK from their technical performance to the operational considerations and device purchasing procedures. The report concluded that there was no perfect

TABLE 10
Specimen stability and temperature information from manufacturers

Manufacturer/ distributor	Device	Specimen stability and temperature information
Eiken*/Polymedco†	OC-Auto Micro 80 FOB Test System	Manufacturer states specimens are stable for 15 days at 15°C–30°C and 30 days at 2°C–8°C.
Eiken/Polymedco	OC-Sensor DIANA IFOB Test System	There are also data that show the specimen can be kept for less than 3 days at 29°C–34°C but can be kept for at least 10–14 days at −18°C to −24°C
Alfresa Pharma Corporation‡/ Inverness Medical§	I-FOBT Hemoglobin NS-Plus	Manufacturer's marketing materials indicate the specimen is 95% stable for 7 days at 25°C, after 2 days at 37°C stability drops to 90% then to 80% after 7 days, stable for 30 days at −40°C, and after 2 days at 7°C stability drops to 90% but stays at this for 20 days
Beckman Coulter¶	Hemoccult ICT, Immunochemical Fecal Occult Blood Test (also known as Flexsure OBT)	Manufacturer instructions say specimen is stable after sampling for 14 days at 15°C–30°C
Eiken/Polymedco	OC-Light Manual IFOBT	Manufacturer states the specimen is stable for 15 days at 15°C–30°C or 30 days at 2°C–8°C
Inverness Medical	Clearview Ultra FOB Test	Manufacturer states that specimen can be stored at 15°C–30°C for up to 5 days or 2°C–8°C for up to 14 days.
Medix Biochemica**	Actim Fecal Blood Test	Manufacturer states that specimen is stable for up to 7 days at 2°C–25°C
PSS World Medical††	Consult Diagnostic Occult Blood Test Extra Sensitive	Unknown.
Artron‡‡	One Step Fecal Occult Blood Test	Manufacturer instructions state the test should be developed immediately and read within 10 min to 15 min. No information on storage if not developed immediately.
IND Diagnostic§§/ BTNX¶¶	Rapid Response One-Step Fecal Occult Blood Test	Manufacturer instructions state if not developed straight away the specimen is stable up to 7 days at 37°C. This is intended to be a physician-developed test (although not licensed for this currently) but is suitable and licensed for laboratory development
Tremblay Harrison***	Minute Lab Fecal Occult Blood Test Device	Manufacturer instructions intend for the test to be developed within 6 h of collecting specimen; if not developed within 6 h, specimen is stable at 2°C–8°C for 3 days
WHPM Bioresearch & Technology†††	Hemosure Immunological Fecal Occult Blood Test	Manufacturer instructions intend for the test to be developed by the patient immediately but if not the specimen is stable at 2°C–8°C but they do not state for how long
Innovacon‡‡‡	FOB One Step Fecal Occult Blood Test	Manufacturer instructions intend for the test to be developed by the patient within 1 h, but if not it will be stable for 3 days at 15°C–30°C

**Eiken Chemical Company Ltd, Japan; †Polymedco Inc, USA; ‡*Alfresa Pharma *Corporation, Japan; §Inverness Medical, Canada; ¶Beckman Coulter Inc, USA; **Medix Biochemica, Finland; ††PSS World Medical Inc, USA; ‡‡Artron Bioresearch, Canada; §§IND Diagnostic Inc, Canada; BNTX Inc, Canada; ***Tremblay Harrison Inc, USA; †††WHPM Bioresearch & Technology Inc, USA; ‡‡‡Innovacon Inc, USA. FOB Fecal occult blood*

FIT on the market but that the OC-Sensor DIANA analyzer, despite not being an ideal test, was the most suitable system for the English Bowel Cancer Screening Program.

The Health Council of the Netherlands report on "A national colorectal cancer screening programme" (26) recommended establishing a nationwide CRC screening program using FIT for 55- to 75-year-olds on a biennial basis. The report stated that 50 ng/mL was the optimum positivity threshold in hemoglobin concentration in terms of cost effectiveness, but provisionally recommended a cut-off level of 75 ng/mL because of considerations of colonoscopy capacity required to support the program. A single sampling method was advised due to concerns that increasing sensitivity through multiple sampling may result in decreased participation. The report also recommended that laboratory analysis be organized so that samples could be tested as soon as possible following arrival of the kit and, that when rapid testing is not possible, the sample should be placed in cold storage.

The European Guidelines for Quality Assurance in Colorectal Cancer Screening and Diagnosis (25) state that FITs provide an improvement in the test characteristics over gFOBT due to improved sensitivity and specificity, as well as the ability to automate test development and adjust the concentration at which a positive result is reported. The European Union (EU) Guidelines state that, although FITs are currently the test of choice for population screening, individual device characteristics, including ease of use by the participant and the laboratory, suitability for transport, sampling reproducibility ,and sample stability are important when selecting the FIT device most appropriate to a specific screening program. The EU Guidelines recommend that, until more stability data are published on FIT, screening programs should adopt the conditions and period of storage described in the manufacturer's instructions for use after having determined that they are appropriate for local conditions. They also recommend that consideration should be given to using more than one specimen together with criteria for assigning positivity that, combined, provide a referral rate that is clinically, logistically and financially appropriate to the screening program. The EU Guidelines state that the proportion of unacceptable tests received in the laboratory should not exceed 3% of all kits received and that less than 1% is desirable. The Guidelines note that the proportion of unacceptable tests is influenced by the ease of use of the test kit and the quality of the test kit instructions for use. They recommend that the laboratory be able to unambiguously identify the subject identification on the test device, possibly through the use of barcodes. In addition, the EU Guidelines recommend that a local pilot study should be undertaken to ensure that the chosen device and associated distribution, sampling, and labelling procedures are acceptable.The recommendations are summarized in Table 11.

Implementing FIT in population-based CRC screening programs in Canada: *Nova Scotia:* In 2009, Nova Scotia launched a biennial screening program for men and women 50 to 74 years of age using Hemoccult ICT (Beckman Coulter, USA), a card-sampling method with a non-numeric result. Hemoccult ICT was chosen due to cost considerations, its long shelf-life and its ability to withstand temperature fluctuations. Currently the program is implemented in all district health authorities in Nova Scotia. Participants are invited by letter to participate and receive a FIT kit by mail two weeks later. They complete two samples over two days and return the completed FIT kits by regular post to a single central laboratory. FIT kits with samples older than 10 days and kits with no sample collection date are not processed; in these cases, the participant is sent a second FIT kit for completion, along with a letter explaining why the test could not be processed. An evaluation of the program is underway.

Saskatchewan: In 2009, Saskatchewan launched a pilot program in the Five Hills Health Region for men and women between 50 and 74 years of age; an expansion to cover approximately one-half of the province

TABLE 11
Recommendations for use of fecal immunochemical tests (FIT) in organized colorectal cancer screening programs

Guideline document/author (reference), year	Recommendations
Levin et al (27), 2008	Annual testing with either gFOBT or FIT in both male and female participants 50 years of age or older are both acceptable options for colorectal cancer screening.
Centre for Evidence-Based Practice in the NHS (24), 2009	Compared: Hem-SP/MagStream HT*, OC-Sensor†, FOB, Gold/SENTiFOB‡, FOB Gold DEVEL-A-TAB‡ OC-Sensor DIANA analyzer, despite not being an ideal test, most suitable for the English Bowel Cancer Screening Program.
Health Council of The Netherlands (26), 2009	Nationwide screening program using FIT for 55- to 74-year-olds biennially 50 ng/mL is the optimum cut-off level in hemoglobin concentration in terms of cost-effectiveness. Provisionally recommend cut-off level in hemoglobin concentration of 75 ng/mL due to colonoscopy capacity. Single sampling method advised to maximize positivity Samples should be tested as soon as possible once returned to the laboratory.
Rex et al (28), 2009	Annual FIT testing is the preferred colorectal cancer screening method compared with gFOBT.
European Guidelines for Quality Assurance in Colorectal Cancer Screening and Diagnosis (25), 2010	FIT is preferred over gFOBT for population screening for colorectal cancer. FIT factors such as ease of use, transportability, sample reproducibility, and sample stability need to be considered when developing a program. Prior to implementation, a pilot study should be performed to ensure that the FIT program chosen achieves a positivity rate that is clinically acceptable, logistically and financially possible. The acceptable loss of completed tests is less than 3% of all tests, with the goal being less than 1% Subject IDs should be easily identifiable, possibly through the use of barcodes. Screening programs should adopt manufacturer's storage conditions. A local pilot study of FIT should be conducted before widespread implementation.

**Fujirebio Inc, Japan and Fujirebio Diagnostics Inc, USA; †Eiken Chemical Company Ltd, Japan; ‡Sentinel Diagnostics, Italy; gFOBT Guaiac fecal occult blood test; ID Identification; NHS National Health Service*

TABLE 12
Use of fecal immunochemical test (FIT) in colorectal cancer screening programs in Canada

Province	Stage of implementation	FIT used	Details of sampling and transportation	Details of processing
Nova Scotia	Program is implemented across the province	Hemoccult ICT* (card method)	Two samples over two days. Completed kits returned by business reply mail to a single central laboratory	Samples older than 10 days or with no collection date given are not processed – a letter with explanation along with a second kit are sent to the participant
Saskatchewan	In pilot phase	OC-Auto Micro 80† (vial method)	Completed kits are dropped off at a medical laboratory or mailed at a Canada Post retail outlet in a supplied postage-paid envelope	If the sample cannot be dropped off within 24 h, participants advised to refrigerate sample
British Columbia	In pilot phase	OC-Auto Micro 80 (vial method)	Two samples over two days. Completed kits are dropped off at designated locations, couriered to a central laboratory for processing	Participants advised to refrigerate sample but not to freeze them until returned. Samples are rejected if received more than 15 days after first sample taken

**Beckman Coulter Inc, USA; †Eiken Chemical Company, Ltd, Japan*

was implemented in 2011. The program uses the OC-Auto Micro 80 device (Eiken Chemical Company, Ltd, Japan), a vial-sampling method. Participants are invited by letter to participate and receive a FIT kit by mail three weeks later, with reminders sent after six weeks if the kit has not been returned. Kits are barcoded with the name, date of birth and sex of the participant; the participant also completes a form that verifies his or her eligibility for screening and identifies the primary care provider to be notified of test results. Eligible (ie, average risk) participants complete one sample and return the completed FIT kit to a designated medical laboratory or a Canada Post retail outlet in a supplied postage-paid envelope. Completed kits cannot be sent by regular post. If the sample cannot be dropped off within 24 h, participants are advised to refrigerate the sample. Approximately 3.5% of kits have been rejected, largely due to the specimen collection date not being marked, expired samples and damaged kits. An evaluation of the pilot program is underway.

British Columbia: In 2009, British Columbia launched a pilot program in Penticton, Powell River and the downtown core of Vancouver for men and women 50 to 74 years of age, using the OC-Auto Micro 80 device, a vial-sampling method. The participants request a kit through a toll-free number and collect two samples over two days. Participants are advised that the completed kit must be stored in the refrigerator, and not frozen, and returned to a specified drop-off location (designated laboratory, hospital or physician's office), ideally within one or two days; from there, kits are sent to a central processing laboratory in Vancouver by courier. Completed kits cannot be sent by regular post. Results must be analyzed within 15 days after the first sample is collected. Table 12 summarizes the details of FIT use in Canada.

DISCUSSION

The purpose of the present evidentiary review was to evaluate the existing evidence concerning FIT to inform the decision on how to replace the current gFOBT with FIT in Ontario's ColonCancerCheck Program.

FIT performance factors

At this time, and in contrast to gFOBT, there is no evidence from RCTs involving average-risk screening populations concerning the use of FIT in repeated (annual or biennial) testing. However, gFOBT used in repeated testing (coupled with a colonoscopy for those who screen positive) is associated with a reduction in CRC mortality (4-6). Therefore, the published evidence evaluated here compared the test characteristics of FIT with gFOBT in one-time (not repeated) testing. The assumption is that, if the test characteristics compare favourably, FIT used in a screening program with repeated testing would, at a minimum, achieve the same mortality reduction. A detailed evaluation of the test characteristics of FIT for detecting AA is beyond the scope of the present work. However, to the extent that FITs are able to detect AA, the use of FIT for CRC screening holds promise for CRC prevention as well as early detection.

Results from two Dutch RCTs (17,19) involving asymptomatic persons at average risk for CRC show that, compared with a standard gFOBT, the CRC and AA detection rates are greater with FIT but that specificity is lower. The PPVs of FIT and gFOBT for detecting CRC and AA are similar, but the positivity rate of FIT (when used according to the cut-off level in hemoglobin concentration recommended by the manufacturer to define a positive test) is greater. Note, however, that the positivity rates reported, which were 5.5% and

4.8% in the two Dutch RCTs, are comparable with the current positivity rate of 4.7% observed in the ColonCancerCheck Program. The increase in CRC detection is an advantage of FIT compared with gFOBT. In addition, the increased detection of AA is a key advantage of FIT compared with the standard gFOBT, which does not detect AA. Because AAs are precancerous lesions, the detection of AA (and their removal at colonoscopy) means that FIT use may be associated with CRC prevention. In the Dutch RCTs, only individuals who had a positive FIT underwent colonoscopy; therefore, sensitivity could not be evaluated. In the study of Allison et al (20), the investigators compared the sensitivity of gFOBT and FIT for the detection of CRC in the distal colon. In this study, the sensitivity of FIT for detecting CRC was greater than for gFOBT, although the difference was not statistically significant. However, because Allison et al used a sensitive gFOBT (HemeSENSA), and not a standard gFOBT, the results cannot be generalized to Ontario, where the ColonCancerCheck program uses a standard gFOBT. The RCT of Park et al (23) reported a superior sensitivity for FIT compared with a standard gFOBT for detecting CRC and AA.

FIT positivity rates are affected by the number of stool specimens sampled and the definition of a positive test. One study compared FIT performance for single versus multiple stool samples and reported that a one-sample method resulted in higher positivity than a two-sample method in which both tests had to provide a positive result to be considered positive. A two-sample method, in which only one test had to provide a positive result to be considered positive, resulted in the highest positivity rate. The recommendations of kit manufacturers vary, but most manufacturers advise a single stool sample.

The two Dutch RCTs (17,19) used FIT kits that provide a numerical result for hemoglobin concentration and, in additional studies (16,18), reported on the effects of a change in cut-off level in hemoglobin concentration on FIT performance. These results showed that when lower cut-offs in hemoglobin concentration were used to define a positive FIT, the detection of CRC and AA were greater, the specificity and PPV for CRC lower, and the positivity rates were higher. The only study that reported on sensitivity at different cut-off levels in hemoglobin concentration was that by Park et al (23), who reported that sensitivities for CRC and AA were decreased as the cut-off level increased. Based on their results, both groups of Dutch investigators recommended using a cut-off in hemoglobin concentration below the manufacturer's standard cut-off of 100 ng/mL, with one group (18) advising a cut-off of 75 ng/mL. Choosing an optimal cut-off in stool hemoglobin concentration for screening an average-risk population involves weighing the better clinical outcomes (more CRCs and AA detected) associated with lower cut-offs against the higher costs (more colonoscopies required).

All three RCTs reported statistically significantly higher participation rates for FIT compared with gFOBT. The Expert Panel noted that the ColonCancerCheck Program needs to maximize screening participation rates. An evaluation of the factors that affect participation rates was beyond the scope of the present review. However, explanations for higher participation rates are linked to kit usability factors, which are discussed below.

In summary, with respect to FIT performance, compared with standard gFOBT, current evidence indicates an increased participation in screening, higher sensitivity for the detection of CRC and AA, and higher rates of detection for CRC and AA. It is important to recognize that detecting AA is a distinguishing feature of FITs compared with standard gFOBTs. These advantages of FIT are offset by a lower specificity for the detection of CRC and AA, and a higher positivity rate (when using the manufacturer's cut-off), which in turn may require a greater number of colonoscopies. However, these performance characteristics change when the cut-off level in hemoglobin concentration is changed, allowing a screening program to select the optimal cut-off for the program, balancing the better clinical outcomes (more cancers detected) associated with lower cut-offs against the higher costs (more colonoscopies required).

FIT kit usability factors

Because the FIT is specific for human hemoglobin, there is no interference from dietary substances and, in general, dietary restriction is not advised. No published studies evaluating diet or medication use were identified, although three manufacturers of FIT kits that do not provide a numerical result advised avoidance of alcohol and acetylsalicylic acid and similar medications for 48 h before sample collection. In terms of dietary and medication restrictions, the ColonCancerCheck Program advises only that vitamin C supplements and citrus fruit and juices be avoided for three days before and during stool sample collection with gFOBT (9).

Specimen stability

While sample stability has not been a major issue for gFOBT, it is a consideration with FIT because of the relative instability of the globin in the collection systems used. Temperature, which can be affected by weather, transport and storage conditions, affects specimen stability. The Italian study by Grazzini et al (21) showed that during the warmer summer months, the test characteristics of FIT differed from those in the winter months. The authors reported that a 1°C increase in temperature reduced the probability of a positive FIT by 0.7% and that this resulted in a 13% reduction in the probability of detecting a CRC or AA in summer compared with winter. Because of this, the manufacturers specify storage and transport conditions to minimize the effect of sample instability on FIT performance. In general, compared with gFOBT, the conditions for FIT are more stringent and the time period between sample collection and processing is shorter. Satisfying these more stringent conditions is challenging for organized CRC screening programs.

In summary, specimen stability with regard to temperature and time is an issue that requires consideration and a thorough understanding of the specifications of the FIT device chosen. The Expert Panel discussed the issue of specimen stability given the extreme temperatures that occur in Ontario, and recommend that a pilot study be conducted before full implementation, in part to assess specimen stability (see below).

Lessons from ColonCancerCheck

Ontario's ColonCancerCheck Program was launched province-wide in 2008. The Program uses a standard (not high sensitivity) gFOBT (Hemascreen [Immunostics Inc, USA]), which has the same test characteristics as the Hemoccult II. This is the same gFOBT currently used in the English Bowel Cancer Screening Program. Participants obtain a gFOBT kit from their primary care provider or, if they do not have one, from a community pharmacy or by calling Telehealth Ontario. Kits are not mailed to participants. Two samples are collected from each of three stools. Participants can return completed kits by regular mail or by drop-off at a participating laboratory. The majority of kits are returned by mail. A positive test is defined as at least one positive sample. The positivity rate was relatively stable throughout 2010 (approximately 4.7%).

The experience gained over the past two years highlights the importance of specific aspects of program design that are independent of the type of kit used (ie, gFOBT or FIT). Understanding these aspects provides an opportunity for improving the performance of the ColonCancerCheck Program.

According to Ontario regulations, for a laboratory to process a test, the test must be accompanied by a completed, signed requisition form, and both the test and the requisition must contain two matching unique patient identifiers: typically name, date of birth and/or Ontario Health Insurance Plan Number. If these conditions are not met, the test will be rejected and not processed by the laboratory. gFOBT kits are also rejected if the specimens are more than 21 days old (9). gFOBT kits are considered to have an indeterminate result if a window has a negative result but no specimen collection date marked, or if the sample was applied incorrectly to a window. Participants whose kits are rejected for processing or yield an indeterminate result are advised to

obtain another kit and repeat the test. The percentage of gFOBT kits rejected for processing has declined from 16% at program launch (2008) to 4% in 2010; still much higher than the EU Guideline recommendation of less than 1% (25). The percentage of kits with indeterminate results was stable in 2010 (6%). Taken together, more than 10% of participants are advised to repeat the test. Many of these initial participants do not subsequently submit a satisfactory sample. This clearly represents a missed opportunity to detect CRC and consumes considerable resources.

The current unacceptably high rate of rejected specimens and indeterminate results is largely due to program design, particularly the way in which the gFOBT kits are distributed and labelled. Family physicians and/or patients complete the required information (name and date of birth) on both the kit and the requisition; regulations require that this information must match exactly before a laboratory can process a kit. Prelabelling the kits with unique patient identifiers and eliminating the need for a separate requisition would dramatically reduce the unacceptably high rate of rejected kits and ensure the improved performance of the ColonCancerCheck Program and a better use of resources.

Implementing FIT in population-based CRC screening programs

In Europe, guidelines have made recommendations for FIT device selection and implementation in population-based CRC screening programs (24-26). Recognizing the potential challenges of launching a FIT-based CRC screening program, the guidelines have recommended pilot programs to ensure that all logistical challenges are addressed before full implementation (25,26).

One Canadian province (Nova Scotia) has achieved province-wide program implementation using a card-based FIT, which is associated with greater specimen stability and is returnable by regular post. In contrast, Saskatchewan and British Columbia are piloting the use of vial-based FITs that cannot be returned by mail because of specimen stability concerns.

Recommendations for Ontario's ColonCancerCheck Program

The Expert Panel concludes that the FIT has the following important advantages compared with the standard gFOBT: higher screening participation rates, greater sensitivity for detecting CRC and AA, potential for automation in the laboratory and potential to select the cut-off level of hemoglobin concentration that defines a positive test. However, there are the following potential disadvantages: greater specimen instability and possibly higher positivity rates.

The Expert Panel concludes that the ideal FIT would have the following features:

1. Provide a numerical result (so the cut-off level in hemoglobin concentration can be chosen).
2. Be readily automated in the laboratory.
3. Require one stool sample.
4. Have specimen stability across wide variations in temperature.
5. Have specimen stability for at least seven days between the time of sample collection and processing in the laboratory.

Currently, it is uncertain whether any FIT available in Canada has all of these features. The Expert Panel recommends that Ontario's ColonCancerCheck Program conduct a pilot study to evaluate the performance of one or more FIT kits to guide the selection of a FIT device as well as guide any changes in program design required for FIT implementation. Regarding the investigation of various cut-off values, the Expert Panel believes that this is beyond the scope of the planned pilot study. The pilot study would evaluate the FIT kits in the laboratory and in the field. The laboratory component would include an evaluation of specimen stability under varying conditions and the feasibility of using automated processes in a population-based program. The field component would evaluate kit distribution, labelling of kits, stool sampling, and transportation of completed kits to the laboratory. An economic evaluation should also be conducted. The intent would be to evaluate these aspects in such a way that when the laboratory and field components are combined, the redesigned program would ensure feasibility and improved performance at an acceptable cost.

Finally, based on findings from the current ColonCancerCheck Program, the Expert Panel strongly recommends changes in program design such that the current approach of manual kit labelling be changed to an automated approach (eg, using a barcode), and the need for a separate requisition to accompany the kit be dropped. In this way, the performance of Ontario's ColonCancerCheck Program would be improved, and better use would be made of current resources.

DISCLOSURES: The authors have no financial disclosures or conflicts of interest to declare.

FUNDING: The PEBC is a provincial initiative of Cancer Care Ontario. It is supported by the Ontario Ministry of Health and Long-Term Care through Cancer Care Ontario. All work produced by the PEBC is editorially independent from its funding source.

APPENDIX A
FIT GUIDELINES EXPERT PANEL

Linda Rabeneck MD MPH FRCPC (Chair), Vice President, Prevention and Cancer Control, Cancer Care Ontario, 620 University Avenue Toronto, Ontario M5G 2L7.

R Bryan Rumble MSc, Research Coordinator, Cancer Care Ontario's Program in Evidence-based Care, Department of Oncology, McMaster University, Henderson Site, 60 (G) Wing, 2nd Floor, Room 226, 711 Concession Street, Hamilton, Ontario L8V 1C3.

Frank Thompson MD, 333 Aberdeen Boulevard, Midland, Ontario L4R 5M9.

Michael Mills MD CCFP FCFP, Regional Primary Care Lead, Hamilton Niagara Haldimand Brant LHIN, Cancer Care Ontario, Juravinski Regional Cancer Centre, 699 Concession Street, Hamilton, Ontario L8V 5C2.

Curtis Oleschuk MD FCACB, Clinical Biochemist, Department of Clinical Biochemistry and Genetics, Diagnostic Services of Manitoba MS543, 820 Sherbrook Street, Winnipeg, Manitoba R3A 1R9.

Alexandra Whibley BSc(Hons), Project Coordinator, Cancer Care Ontario, 505 University Avenue, Toronto, Ontario M5G 1X3.

Hans Messersmith MPH, Assistant Director, Quality and Methods, Program in Evidence-Based Care, Cancer Care Ontario, McMaster University, 711 Concession Street, Hamilton, Ontario L8V 1C3

Nancy Lewis PhD, Senior Policy and Planning Officer, Cancer Care Ontario, 505 University Avenue, Toronto, Ontario M5G 1X3.

APPENDIX B
Fecal immunochemical kits approved by Health Canada

Manufacturer /distributor	Device	Class	Product description	Sample, Vial/card	Number of samples	Positive cut-off point	Throughput/ development time	Specimen stability and temperature information	Notes
Eiken/ Polymedco	OC-Auto Micro 80 FOB Test System	2	Flat tube, dipstick collection, machine developed	Vial	1 sample across 1 day although some groups use 2 samples across 2 days	Can be set by user. Machine comes set at 100 ng/mL	80 samples per hour	Manufacturer states samples are stable for 15 days at 15˚C–30˚C and 30 days at 2˚C–8˚C. There is also data that shows the sample can be kept for less than 3 days at 29˚C–34˚C but can be kept for at least 10–14 days at –18˚C to –24˚C[1].	Product insert and patient instructions obtained Being used by British Columbia and Saskatchewan in pilot testing
Eiken/ Polymedco	OC-Sensor DIANA IFOB Test System	2	Flat tube, dipstick collection, machine developed	Vial	1 sample across 1 day although some groups use 2 samples across 2 days	Can be set by user. Machine comes set at 100 ng/mL	280 samples per hour		Used OC-Auto product insert and patient instructions
Alfresa Pharma Corp / Inverness Medical	I-FOBT Hemoglobin NS-Plus	2	Flat tube, dipstick collection, machine developed	Vial	2 samples across 2 days	Set by user	300 samples per hour	Manufacturer's marketing materials indicate the sample is 95% stable for 7 days at 25˚C, after 2 days at 37˚C stability drops to 90% then to 80% after 7 days, stable for 30 days at –40˚C, and after 2 days at 7˚C stability drops to 90% but stays at this for 20 days.	Product insert and marketing slideshow obtained
Beckman Coulter	Hemoccult ICT, Immunochemical Fecal Occult Blood Test (also known as Flexsure OBT)	2	Test card, applicator stick, on card developed	Card	3 samples across 3 days	Unknown	2 min per test	Manufacturer instructions say test is stable after sampling for 14 days at 15˚C–30˚C	Product insert and patient instructions obtained Being used by Nova Scotia in pilot testing
Eiken/ Polymedco	OC-Light Manual IFOBT	2	Long cylindrical tube, dipstick collection, test strip developed	Vial	1 sample over 1 day	Unknown	5 min per test	Polymedco states the specimen is stable for 15 days at 15˚C–30˚C or 30 days at 2˚C–8˚C	No product insert obtained, information taken from www.ifobt.com/hp_overmanualhtml
Inverness Medical	Clearview Ultra FOB Test	2	Long cylindrical tube, dipstick collection, test strip developed	Vial	1 sample over 1 day	50 ng/mL	5 min per test	Manufacturer states that specimen can be stored at 15˚C–30˚C for up to 5 days or 2˚C–8˚C for up to 14 days	Product insert obtained
Medix Biochemica	Actim Fecal Blood Test	2	Cylindrical tube, sampling stick that then is put into the tube, development occurs on the stick	Vial	Unclear but appears to be 1 sample from 1 day	Unknown	10 min per test	Manufacturer states that specimen is stable for up to 7 days at 2˚C–25˚C	No product insert obtained, information taken from www.bhr.co.uk/actim-fecal-blood-test-procedure-2069-0.html
PSS World Medical	Consult Diagnostic Occult Blood Test Extra Sensitive	2	Unknown	Vial	Unknown	Unknown	Unknown	Unknown	Have not been able to gather information on this test
Artron	One Step Fecal Occult Blood Test	2	Cylindrical tube, dipstick sampling, developed on a cassette	Vial	1 sample from 1 day	50 ng/mL	10–15 min per test	Manufacturer instructions state the test should be developed immediately and read within 10–15 min. No information on storage if not developed immediately	Product insert obtained

APPENDIX B - CONTINUED
Fecal immunochemical kits approved by Health Canada

Manufacturer /distributor	Device	Class	Product description	Sample, Vial/card	Number of samples	Positive cut-off point	Throughput/ development time	Specimen stability and temperature information	Notes
IND Diagnostic / BTNX	Rapid Response One-Step Fecal Occult Blood Test	2 (applying for class 3)	Cylindrical tube, dipstick sampling, developed on a cassette	Vial	1 sample from 1 day	50 ng/mL	5 min per test	Manufacturer instructions state if not developed straight away the specimen is stable up to 7 days at 37˚C. This is intended to be a physician developed test (although not licensed for this currently) but is suitable and licensed for laboratory development.	Product insert obtained
Tremblay-Harrison	Minute Lab Fecal Occult Blood Test Device	3	Cylindrical tube, dipstick sampling, developed on a cassette	Vial	1 sample from 1 day	50 ng/mL	5 min per test	Manufacturer instructions intend for the test to be developed within 6 h of collecting sample, if not developed within 6 h, sample is stable at 2˚C–8˚C for 3 days .	Product insert obtained
WHPM Bioresearch & Technology	Hemosure Immunological Fecal Occult Blood Test	2	Cylindrical tube, dipstick sampling, developed on a cassette	Vial	1 sample from 1 day	50 ng/mL	5 min per test	Manufacturer instructions intend for the test to be developed by the patient immediately but if not the specimen is stable at 2˚C–8˚C but they do not state for how long	Product insert obtained
Innovacon	FOB One Step Fecal Occult Blood Test	2	Cylindrical tube, dipstick sampling, developed on a cassette	Vial	1 sample from 1 day	50 ng/mL	10 min per test	Manufacturer instructions intend for the test to be developed by the patient within an hour, but if not it will be stable for 3 days at 15˚C–30˚C	Product insert obtained

Device classification 2 means the product is licensed for development in a laboratory setting only, although Health Canada do not regulate this, physicians could develop the test in their office. Device classification 3 means the product is licensed for development at any point of care, which could be physician's office or pharmacy. Reference: 1 NHS. Evaluation report: Immunochemical faecal occult blood tests. November 2009. See Table 1 for manufacturers and country

APPENDIX C. Definition of diagnostic parameters
Relationship between screening test result and presence of cancer

	Cancer present	
Screening test result	**Yes**	**No**
Positive	True positive (a)	False positive (b)
Negative	False negative (c)	True negative (d)

The definitions used in the present guideline are as follows:

True positive (TP) — Those with a positive screening test and confirmed cancer (a)

False positive (FP) — Those with a positive screening test and no confirmed cancer (b)

True negative (TN) — Those with a negative screening test and no confirmed cancer (d)

False negative (FN) — Those with a negative screening test and confirmed cancer (c)

Positive predictive value (PPV) — Proportion of people with a positive screening test who have confirmed cancer (a/[a+b])

Sensitivity — Proportion of people with cancer who have a positive screening test (a/[a+c])

Specificity — Proportion of people who do not have cancer who have a negative screening test (d/[b+d])

APPENDIX D. Literature search strategies

Database: Ovid MEDLINE(R) <1996 to June Week 2 2010>

1 fecal immunohistochemical test.mp. (0)
2 exp Immunohistochemistry/ or fecal immunochemical test.mp. (263814)
3 screening.mp. or exp Mass Screening/ (182626)
4 colorectal neoplasms.mp. or exp Colorectal Neoplasms/ (67158)
5 2 and 4 (5545)
6 3 and 5 (244)
7 limit 6 to (english language and humans) (227)
8 from 7 keep 1,4,6,10-11,19-20,22,33,35,37-38,40,52,54,66,71-72,76,84,90,95,103,113,115,119,162,184-185,191,218-219,224 (33)

Database: EMBASE <1996 to 2010 Week 23>

1 fecal immunohistochemical test.mp. (0)
2 exp immunohistochemistry/ or immunohistochemicalmp. (194893)
3 fecal immunochemical test.mp. (14)
4 exp screening/ or exp cancer screening/ or screening.mp. or screening test/ (232254)
5 colon cancer.mp. or exp colon cancer/ (73399)
6 rectal cancer.mp. or exp rectum cancer/ (53138)
7 5 or 6 (81276)
8 2 or 3 (194906)
9 4 and 8 (3172)
10 7 and 8 and 9 (385)
11 limit 10 to english language (362)
12 from 11 keep 5,7,12,16,23,29,33,41-42,46,51,53,59,72,84-86,91,96,107,136,143,161,176,214,244,263,266,301,327,362 (31)

APPENDIX E
QUADAS results

	Allison et al, 2007			Park et al, 2010		
Domain	**Yes**	**No**	**Unclear**	**Yes**	**No**	**Unclear**
1. Was the spectrum of patients representative of the patients who will receive the test in practice?	X			X		
2. Were selection criteria clearly described?	X			X		
3. Is the reference standard likely to correctly classify the target condition?		X		X		
4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	X			X		
5. Did the whole sample or a random selection of the sample receive verification using a reference standard of diagnosis?		X		X		
6. Did patients receive the same reference standard regardless of the index test result?		X		X		
7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	X			X		
8. Was the execution of the index test described in sufficient detail to permit replication of the test?	X			X		
9. Was the execution of the reference standard described in sufficient detail to permit its replication?		X		X		
10. Were the index test results interpreted without knowledge of the results of the reference standard?			X			X
11. Were the reference standard results interpreted without knowledge of the results of the index test?			X			X
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	X					X
13. Were uninterpretable/ intermediate test results reported?	X			X		
14. Were withdrawals from the study explained?	X			X		
TOTALS	**8**	**4**	**2**	**11**	**0**	**3**

REFERENCES

1. Canadian Cancer Society's Steering Committee on Cancer Statistics. Canadian Cancer Statistics 2011. Toronto, Ontario, 2011. <www.cancer.ca/canada-wide/about%20cancer/cancer%20statistics.aspx> (Accessed May 2011).
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer 2010;127:2893-917.
3. Canadian Task Force on Preventive Health Care. Colorectal cancer screening: Recommendation statement from the Canadian Task Force on preventive health care. CMAJ 2001;165:206-8.
4. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;348:1472-7.
5. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;348:1467-71.
6. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. N Engl J Med 1993;328:1365-71.
7. Atkin WS, Edwards R, Kralj-Hans I, et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: A multicentre randomised controlled trial. Lancet 2010;375:1624-33.
8. Rabeneck L, Rumble RB, Axler J, et al. Cancer Care Ontario Colonoscopy Standards: Standards and evidentiary base. Can J Gastroenterol 2007;21(Suppl D):5D-24D.
9. Rabeneck L, Zwaal C, Goodman JH, Mai V, Zamkanei M. Cancer Care Ontario guaiac fecal occult blood test (FOBT) laboratory standards: Evidentiary base and recommendations. Clin Biochem 2008;41:1289-305.
10. Young GP, Cole S. New stool screening tests for colorectal cancer. Digestion 2007;76:26-33.

11. Young GP, St. John DJB. Faecal occult blood tests: Choice, usage and clinical applications. Clin Biochem Rev 1992;13:161-7.
12. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: A tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 2003;3:25.
13. Mujoomdar M, Cimon K, Spry C. Fecal immunochemical tests for colorectal cancer screening: A systematic review of accuracy and compliance. Ottawa: Canadian Agency for Drugs and Technologies in Health, 2009:1-13.
14. Whitlock EP, Lin JS, Liles E, Beil TL, Fu R. Screening for colorectal cancer: A targeted, updated systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2008;149:638-58.
15. Hoffman RM, Steel S, Yee EF, Massie L, Schrader RM, Murata GH. Colorectal cancer screening adherence is higher with fecal immunochemical tests than guaiac-based fecal occult blood tests: A randomized, controlled trial. Prev Med 2010;50:297-9.
16. Hol L, van Leerdam ME, van Ballegooijen M, et al. Screening for colorectal cancer: Randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut 2010;59:62-8.
17. Hol L, Wilschut JA, Van Ballegooijen M, et al. Screening for colorectal cancer: Random comparison of guaiac and immunochemical faecal occult blood testing at different cut-off levels. Br J Cancer 2009;100:1103-10.
18. van Rossum LG, van Rijn AF, Laheij RJ, et al. Cutoff value determines the performance of a semi-quantitative immunochemical faecal occult blood test in a colorectal cancer screening programme. Br J Cancer 2009;101:1274-81.
19. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology 2008;135:82-90.
20. Allison JE, Sakoda LC, Levin TR, et al. Screening for colorectal neoplasms with new fecal occult blood tests: Update on performance characteristics. J Natl Cancer Inst 2007;99:1462-70.
21. Grazzini G, Ventura L, Zappa M, et al. Influence of seasonal variations in ambient temperatures on performance of immunochemical faecal occult blood test for colorectal cancer screening: Observational study from the Florence district. Gut 2010;59:1511-5.
22. Grazzini G, Visioli CB, Zorzi M, et al. Immunochemical faecal occult blood test: Number of samples and positivity cutoff. What is the best strategy for colorectal cancer screening? Br J Cancer 2009;100:259-65.
23. Park DI, Ryu S, Kim YH, et al. Comparison of guaiac-based and quantitative immunochemical fecal occult blood testing in a population at average risk undergoing colorectal cancer screening. Am J Gastroenterol 2010;105:2017-25.
24. Lamph SA, Bennitt WE, Brannon CR, Halloran SP. Evaluation Report: Immunochemical Faecal Occult Blood Tests. Guildford: Centre for evidence-based purchasing 2009:4-40.
25. Halloran SP, Launoy G, Zappa M. Faecal occult blood testing In: Segnan N, Patnick J, von Karsa L, eds. European Guidelines for Quality Assurance in Colorectal Cancer Screening and Diagnosis, 1st edn. Luxembourg: Publications Office of the European Union; 2010:103-44.
26. Health Council of the Netherlands. A national colorectal cancer screening programme. The Hague: Health Council of the Netherlands; 2009; publication no. 2009/13E.
27. Levin B, Lieberman DA, McFarland B, et al. Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: A joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. Gastroenterology 2008;134:1570-95.
28. Rex DK, Johnson DA, Anderson JC, et al. American College of Gastroenterology guidelines for colorectal cancer screening 2009. Am J Gastroenterol 2009;104:739-50.
29. Allison JE, Tekawa IS, Ransom LJ, Adrain AL. A comparison of fecal occult-blood tests for colorectal-cancer screening. N Engl J Med 1996;334:155-9.
30. St John DJ, Young GP, Alexeyeff MA, et al. Evaluation of new occult blood tests for detection of colorectal neoplasia. Gastroenterology 1993;104:1661-8.