# San Diego 2016

## *9th Meeting of the Expert Working Group (EWG) – 'FIT for Screening'*

Friday, 20 May 2016: 10:15–12:00

## MEETING REPORT

* * * * *

Expert Working Group (EWG) founding members:

- Jim Allison, University of California, San Francisco, USA (James.Allison@ucsf.edu)
- Callum Fraser, University of Dundee, Scotland (callum.fraser@nhs.net)
- Stephen Halloran, Director of the NHS Bowel Cancer Screening Southern Programme Hub (retired) and Professor Emeritus, University of Surrey, UK (s.halloran@surrey.ac.uk)
- Graeme Young, Flinders University of South Australia, Australia (graeme.young@flinders.edu.au)

The meeting was chaired by Professor Ernst Kuipers, Department of Gastroenterology and Hepatology, Erasmus MC, Rotterdam, The Netherlands (e.j.kuipers@erasmusmc.nl)

Summary report prepared by Helen Seaman (helenseaman@nhs.net)

**Agenda items:**

1. **Welcome & Introduction**
   Ernst Kuipers, Professor of Medicine, Erasmus MC, Rotterdam, NL
2. **One versus two FIT sampling** (slide set no. 1)
   Graeme Young, Flinders Centre for Innovation in Cancer, Flinders University of South Australia
3. **Update from New Zealand** (slide set no. 2)
   Susan Parry, Gastroenterologist & Clinical Director, New Zealand's Ministry of Health Bowel Cancer Programme, NZ
4. **Program sensitivity of FIT applied over time** (slide set no. 3)
   David Ransohoff, Professor of Medicine/Clinical Professor of Epidemiology, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, USA
5. **The impact of FIT positive colonoscopies on the colonoscopy and pathology services** (slide set no. 4). Robert Hilsden, Associate Professor, Departments of Medicine/Community Health Sciences, University of Calgary, AB, Canada
6. **Head-to-head comparisons of different FIT** (slide set no. 5)
   Thomas Imperiale, Professor of Medicine, Indiana University Medical Center, USA
7. **Role of FIT in post-polypectomy surveillance** (slide set no. 6)
   Ann Zauber, Member and Attending, Memorial Sloan Kettering Cancer Center, New York, USA
8. **Colonoscopy versus FIT in reducing mortality from CRC (CONFIRM): impact of mailing time and season on FIT positivity** (slide set no. 7)
   Douglas Robertson and Jason Dominitz, US Department of Veterans Affairs, USA
9. **FIT every year or every other year?** (slide set no. 8)
   T.R. Levin, Clinical Lead, Colorectal Cancer Screening, The Permanente Medical Group, USA

## 1. Introduction

Ernst Kuipers, Department of Gastroenterology & Hepatology, Erasmus MC, Rotterdam, NL

The World Endoscopy Organization's Colorectal Cancer Screening Committee's EWG 'FIT for Screening' was founded in 2012 by Professors Graeme Young (Australia), Stephen Halloran (UK), Callum Fraser (Scotland) and James Allison (USA). The EWG meets twice a year and met for the ninth time in San Diego on Friday 20 May 2016. The meeting attracted more than 80 delegates who heard from a selection of international speakers on new developments in FIT CRC screening around the world. The EWG's remit, publications and meeting reports are available here: http://www.worldendo.org/weo-crcsc-expert-working-group-fit-for-screening.html

## 2. One versus two FIT sampling (slide set no. 1)

Graeme Young, Flinders Centre for Innovation in Cancer, Flinders University of South Australia

Only one FIT sample would be needed if bowel lesions bleed at a consistent rate into a stool of consistent size, but this is not the case as illustrated by the different concentrations of Hb (ng Hb/mL buffer) in two FIT samples for >370 individuals. Second, multiple stool sampling may be more useful when the Hb concentration is close to the cut-off for positivity (*i.e.* multiple stool sampling is less critical for the detection of cancers than for adenomas).

Professor Young described a modelling exercise testing 2-sample FIT prior to colonoscopy based on a screening population of 17,331 individuals at average or increased risk for CRC where colonoscopies have been performed in 2,078 individuals (regardless of the FIT [OC-SENSOR] result). The sensitivity for cancer (23 cases) and number needed to colonoscope (NNC) at different cut-offs for positivity illustrate that if targeting the same sensitivity (a), more colonoscopies were needed using 1-sample FIT, (b) a 2-sample test was more efficient and (c) there was little to be gained by collecting two samples with a cut-off for positivity at 10 µg Hb/g faeces. Similarly, the figures for advanced neoplasia (375 cases) illustrated that (a) 2-sample FIT detects more advanced neoplasia and remains almost as efficient as 1-sample FIT and (b) 2-sample FIT at 20 µg Hb/g faeces is more efficient and more sensitive than 1-sample FIT at 10 µg Hb/g faeces. Professor Young illustrated the observations using a 'DEW' (Detection, Effort and Workload) bubble chart.

The principles emerging from modelling may be summarised thus:

- for high CRC sensitivity ≈90%, two samples at 20 µg Hb/g faeces are slightly more efficient than one sample at 10 µg Hb/g faeces,
- for CRC sensitivity ≈80%, two samples at 40 µg Hb/g faeces are most efficient, and
- for detection of advanced lesions, two samples at 20 µg Hb/g faeces are better than one or two samples at 10 µg Hb/g faeces.

The literature on participation in screening with 1-sample FIT versus 2-sample FIT is conflicting. A study from Australia suggests that participation is less with 2-sample FIT ('Two-stool mode: 199/600, 33.2%; One stool mode: 223/600, 37.2% (p=0.16)' (1)), although a Dutch study found little difference (61.5% in the 1-sample group and 61.3% in the 2-sample group (2)).

A study from French investigators (3) provided a comparison of the performance of three FOBTs under study conditions and following the manufacturers' guidelines: OC-SENSOR (30 µg Hb/g faeces [150 ng Hb/mL buffer)), MagStream (180 µg Hb/g faeces (55 ng Hb/mL buffer) and Hemoccult II (a guaiac-based test). For each FIT, compared with one sample, adding a second sample at the same cut-off value decreased the number needed to screen by increasing the number of colonoscopies but at the expense of an increase in the number needed to scope.

Dutch investigators plotted a ROC curve for positivity versus detection of advanced neoplasia at different cut-off values (2). Per screening strategy, the data points represent the results at cut-off values in the range of 50–200 ng Hb/mL buffer, increasing in steps of 25 ng. For each screening strategy, a higher cut-off level is associated with lower detection, *i.e.* the data points at the left end represent the results at a cut-off value of 200 ng Hb/mL, whereas the data point at the right end represents the results at a cut-off value of 50 ng Hb/mL. The arrows at a positivity of 3.2% and 6.2% define zones in which either 1- or 2-sample FIT screening forms the most efficient strategy.

Professor Young concluded that there is no obvious participatory advantage associated with 1- or 2-FIT sampling, that decisions can be influenced by whether the goal is detecting cancers or advanced neoplasia and that a 3-dimensional DEW analysis shows 2-samples at 20 µg Hb/g faeces is best. Further work should include fine adjustments to the criterion value based on the ROC curve and full cost analysis, including costs of test kits and small differences in participation.

3. **Update from New Zealand** (slide set no. 2)
   Susan Parry, Gastroenterologist & Clinical Director, New Zealand's Ministry of Health Bowel Cancer Programme, NZ

In Professor Parry's absence, her slides were presented by Professor Ernst Kuipers.

A four-year bowel cancer screening pilot commenced in New Zealand in 2012 (4). Men and women aged 50-74 are offered biennial FIT screening using OC-SENSOR (cut-off for positivity 15 µg Hb/g faeces); invitations are generated from a population register. Combined data for participation (defined as return of an 'adequate' kit) in the first and second screening rounds illustrate that participation increases with age, but that participation is poor amongst Pacific Islanders and the most deprived sectors of the population (deprivation deciles 9 & 10). Positivity was 7.3% at the prevalent (first) screen and 5.4% for subsequent screens. Similarly, the detection of, and PPVs for, CRC and adenoma declined from Round 1 to Round 2 (*e.g.* detection of CRC 2.8/1,000 screened in Round 1 versus 1.4/1,000 screened in Round 2). Overall, 192 cancers were detected in Round 1 and 92 in Round 2. Of all cancers detected (as at February 2016), 47.6% were TMN Stage 1 at diagnosis, 21.6% Stage 2, 22.9% Stage 3 and 7.9% Stage 4. The haemoglobin concentration is recorded for all participants allowing assumptions to be made for performance at different cut-off thresholds: as the concentration of Hb increases, positivity decreases, the PPV for CRC and AA increases and the detection of CRC per 1,000 screened declines. The reduction in colonoscopy at 50 µg Hb/g faeces compared with 15 µg Hb/g faeces (the Programme's chosen cut-off for positivity) would have been 50.4% with 79.6% cancers detected.

4. **Program sensitivity of FIT applied over time** (slide set no. 3)
   David Ransohoff, Professor of Medicine/Clinical Professor of Epidemiology, Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, USA

The standard definition of the clinical sensitivity of FIT is the proportion of individuals with disease (true state) that is detected by the test (application sensitivity). The *programme* sensitivity of FIT may be defined as the proportion of individuals with disease that is detected by a programme of repeat testing over time. Can programme sensitivity be estimated knowing the application sensitivity? Assuming FIT results are *independent* of results from previous screening, then programme sensitivity will increase over time. In modelling studies conducted by the USPSTF (United States Preventive Services Task Force), FIT results are assumed to be independent (5).

Professor Ransohoff used the application sensitivity data from the multitarget stool DNA testing for CRC screening paper (6) to illustrate which approach might prove to be more sensitive over time with repeated screens.

The dependence or independence of a screening test might be determined by biology. For example, if lesions never bleed or bleed only at a very late stage, FIT will be less useful. Similarly, if some lesions do not have DNA mutation/methylation, a DNA test will be less useful. But usually we don't know the biology.

Professor Ransohoff referred to the Jensen paper (7) that assessed FIT sensitivity at each application over four years (annual screening). Amongst 323,349 individuals, CRC was diagnosed after a positive FIT (OC FIT-CHEK, Polymedco; cut-off for positivity 20 μg Hb/g faeces) or because of symptoms within a year of screening ('look-back'). Screening detected 80.4% of CRC cases within one year of testing (84.5% in round 1, 73-78% in each subsequent round). Unknowns were the true disease state in all participants, if a longer follow-up would provide different results and whether the stage distribution of disease was different.

Professor Ransohoff then referred to the van der Muelen paper (8) that estimated the proportion of adenomas that do not bleed and may therefore be missed by FIT. MISCAN models were used to fit findings from the Dutch CORERO FIT screening trial (9), using different estimates of test-dependence. The investigators reported that FIT systematically missed about 28% of advanced adenoma (the investigators did not report on CRC).

Professor Ransohoff concluded that programme FIT sensitivity is very difficult to measure or estimate and suggested that in empirical studies longer follow-up and comparisons of stage distribution would be helpful and that modelling studies should include sensitivity analyses.

5. **The impact of FIT positive colonoscopies on the colonoscopy and pathology services** (slide set no. 4). Robert Hilsden, Associate Professor, Departments of Medicine/Community Health Sciences, University of Calgary, AB, Canada

FIT replaced the guaiac-based faecal occult blood test for CRC screening in Alberta, Canada, in November 2013. Invitations for screening using a single FIT (OC-SENSOR) are offered by the Primary Care Physician (PCP) about every two years. The target age range for FIT screening is 50-74 years and the cut-off for positivity is 75 ng Hb/mL buffer (equivalent to 15 μg Hb/g faeces with OC-SENSOR).

The setting for this study was a large non-hospital endoscopy unit in Calgary (six endoscopy suites and about 45 endoscopists) funded for 17,400 colonoscopies each year. Data for 2013 and 2014 were reviewed for all screening-related colonoscopies (primary screening colonoscopy of average or increased risk individuals, surveillance colonoscopies and diagnostic colonoscopies after a positive FIT or CTC). Data on FIT positivity were provided for all ages (<40, 40-49, 50-74 and 75+) – positivity was high for individuals < 40 (11%) and highest for those aged 75+ (16%). After the introduction of FIT at the end of 2013, the volume of gFOBt+ve/FIT+ve referrals for colonoscopy increased steadily over the following year (from about 100 in 2013 to 400-500 per month in 2014) and the volume of average-risk colonoscopy referrals fell from between 1,000 and 1,500 per month in 2013 to about 500 per month in 2014. The detection of screen-relevant CRC lesions or advanced adenoma by colonoscopy amongst FIT+ve individuals was 10 percentage points or more greater than amongst the average-risk individuals. Compared with FIT+ve colonoscopies, a greater proportion of average-risk colonoscopies took $\leq$ 20 minutes (77% vs. 57%), although withdrawal time was longer than 13 minutes for a greater proportion of FIT+ve colonoscopies (45% vs 18%). The use of an EndoClip was markedly higher in FIT+ve colonoscopies (17% vs. 5%) and a greater proportion was referred for surgical follow-up (4% vs. < 1%). In conclusion, introduction of FIT in 2013 had a substantial impact of the colonoscopy service in Calgary, with some of the additional strain caused by inappropriate use of FIT.

**6. Head-to-head comparisons of different FIT** (slide set no. 5)
Thomas Imperiale, Professor of Medicine, Indiana University Medical Center, USA

'What makes a FIT good?' and 'what makes a good FIT?'. Professor Imperiale summarised the features of FIT that illustrate its superiority over gFOBt (*e.g.* specific for human blood, improved uptake) and provided a useful overview of the analytical measures that are used to assess the performance of FIT analysers in the laboratory (*e.g.* analytical sensitivity, precision, linearity and stability) (10).

A UK evaluation of four FIT systems (OC-SENSOR DIANA, Eiken Chemical Co. Ltd.; HM-JACKarc, Kyowa Medex Co. Ltd.; NS-Plus C15 Hb, Alfresa Pharma Corp.; FOB Gold/BioMajesty, Sentinel CH. SpA/Jeol) reported on linearity in the range 0-500 µg Hb/g faeces (slide 7, left-hand side OC-SENSOR DIANA and FOB Gold/BioMajesty) and 0-120 µg/g (slide 7, right-hand side) and described a positive bias for all analysers (10).

French investigators compared three FIT systems (MagStream, OC-SENSOR and FOB Gold) and reported that OC-SENSOR was superior in terms of reproducibility and stability at higher temperatures (11). Another evaluation from the UK assessed the analytical performance of 11 FOBt, including six gFOBt and five immunochemical devices (one quantitative device) (12). The parameters included in the table reproduced on slide 11 are those that showed most variation between the different devices.

A more recent report from the US described an evaluation of 14 FIT devices, all but two of which (both automated FIT) were CLIA-waived, and concluded that many FIT perform acceptably well but others should probably not be used for CRC screening (13). Professor Imperiale then listed the seven FIT currently available in the US, only one of which (FOBT-Chek, Polymedco) is automated.

Comparisons of the clinical performance of different FIT (*i.e*. uptake, sensitivity/specificity *etc*.) can be made using head-to-head comparisons in large populations, modelling (14) and reviews of the literature (2+ FIT and systematic reviews) (15-17). The ideal study would include a population of 50-70 year-olds (± 5 years) at average risk for CRC, all of whom undergo a colonoscopy after completing several different FIT using the same stool sample to identify CRC and AA. Such a study would require a population of between 18 and 25K.

In reality, comparative studies vary with regard to FIT characteristics, performance metrics and colonoscopy (*i.e.* colonoscopy for all regardless of FIT result or colonoscopy/flexible sigmoidoscopy depending on FIT result). Seven studies that compare at least two FIT were listed (3, 18-23) and used to illustrate the various approaches, population sizes and discrepant conclusions about the performance of different FIT.

Professor Imperiale described an on-going meta-analysis of FIT studies by him and colleagues at the Indiana University Medical Center (slide 36) that has included 20 studies and suggested that a network meta-analysis could be useful (24). (A network meta-analysis in this context is a meta-analysis in which multiple treatments (three or more) are compared using both direct comparisons of interventions within randomized controlled trials and indirect comparisons across trials based on a common comparator.)

Professor Imperiale concluded that the published literature is replete with head-to-head comparisons of FIT, although a comprehensive comparison of FIT remains challenging because of variations in study design, test threshold *etc*. The choice of FIT requires consideration of the screening setting, volume of tests anticipated and available resources and FIT programmes require close monitoring to ensure continued performance.

**7. Role of FIT in post-polypectomy surveillance** (slide set no. 6)
Ann Zauber, Member and Attending, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Colonoscopy is currently the recommended surveillance strategy for patients diagnosed with adenomatous polyps. However, a faecal immunochemical test (FIT) could have a role in such surveillance. The National Polyp Study (NPS), a randomized clinical trial that commenced in the 1990s, provides an early example of using faecal occult blood testing (FOBT) for post-polypectomy surveillance. The NPS evaluated the effective surveillance of patients with one or more colorectal adenomas and demonstrated that CRC could be prevented by colonoscopic removal of adenomatous polyps, thereby reducing CRC mortality (25-30). The primary strategy used in the NPS was surveillance colonoscopy. However, the trial also included an annual FOBT with high specificity but relatively low sensitivity (Hemoccult II), which provided little added value to surveillance. Since the NPS, many new FOBTs have been developed and the current FITs, with greatly improved sensitivity for CRC, could now be considered for post-polypectomy surveillance.

In addition, investigators in Australia (31) explored the potential for using FIT in colonoscopic surveillance of patients with a family history or past neoplasia and concluded that an annual FIT detected neoplasia sooner than scheduled surveillances. Subjects with negative FIT had the lowest risk for the most advanced stage of neoplasia and interval FIT analyses could be used to detect missed or rapidly developing lesions in surveillance programmes. Therefore, between FIT's high sensitivity for CRC and the reduced level of risk in those having undergone polypectomy, FIT could be a potentially interesting tool for CRC surveillance.

Dr. Zauber and colleagues have used microsimulation modelling to assess the effects of surveillance colonoscopy. The major impact of colonoscopic polypectomy results from the initial polypectomy, which places the patient at decreased risk for CRC. Subsequent surveillance colonoscopy has less of an effect on CRC mortality reduction than that initial polypectomy. In this way, using FIT for surveillance could be considered clinically reasonable due to the lowered risk of CRC post-polypectomy.

Dr. Zauber and colleagues (Ethna McFerran and Iris Lansdorp-Vogelaar) are currently conducting a microsimulation modelling study of a cohort of 10,000,000 55-75 year-olds using biennial FIT (cut-off 100 ng Hb/mL buffer) with baseline polypectomy followed by FIT surveillance. The model assumes full adherence to screening and standard assumptions on colonoscopy reach, FIT sensitivity and complications from polypectomy (5, 32). Outcome measures will include quality-adjusted life-years (QALYs) gained and CRC mortality. This analysis is in progress; we apologize that our final findings were not available for this meeting. Please contact Dr. Zauber for further details.

**8. Colonoscopy versus FIT in reducing mortality from CRC (CONFIRM): impact of mailing time and season on FIT positivity** (slide set no. 7)
Douglas Robertson and Jason Dominitz, US Department of Veterans Affairs, USA

The CONFIRM study aims to compare outcomes (CRC mortality and incidence) at 10 years amongst 50,000 Veterans aged 50-75 at average-risk for CRC and randomised to either a once-only screening colonoscopy or annual FIT. In February 2016, 35,000 individuals had been enrolled by one of the thirty-six centres participating across the US. One FIT sample is required (OC-FIT CHEC, Polymedco, Inc.) and the cut-off for positivity is > 20 µg Hb/g faeces. In one central laboratory, FIT samples that arrive within 15 days of collection are processed (although undated samples are also processed). FIT samples are refrigerated upon arrival in the laboratory.

Investigators have reported lower faecal Hb concentrations and/or positivity during the summer months in Italy (33), Korea (34), the Netherlands (35) and France (36). Dutch investigators have also reported conflicting data on the association between positivity and increasing transit time (35, 37).

The CONFIRM investigators have assessed the impact of season and transit (shipping) time on FIT positivity. The data were derived from 22,957 processed FIT kits; 12.3% did not have a sample collection date but there was no significant difference in FIT positivity between undated and dated samples (p=0.42). FIT positivity during the summer months (June – August) was 5.8%, compared with between 7.3% and 6.9% during the cooler months (p=0.007). There was no apparent association between FIT positivity and transit time up to 14 days (p=0.56).

The CONFIRM investigators concluded that their findings support processing of undated samples and that improved buffers and/or avoiding mailing during summer months might help to alleviate the threat of seasonal positivity for FIT-based screening programmes. A recent French study from Dancourt *et al* (38) reported no significant variation in FIT positivity according to season or transit time using FOB Gold's new buffer (Sentinel, Milan) (although decreases in the Hb concentration and subsequent positivity data will depend upon the cut-off chosen for positivity).

**9. FIT interval: 1 vs. 2 years** (slide set no. 8)

T.R. Levin, Clinical Lead, Colorectal Cancer Screening, The Permanente Medical Group, USA

The Kaiser Permanente (KP) Medical Care Program in Northern and Southern California serves about 8 million members. Uptake of FIT screening offered using mail outreach has increased since the Program focussed on improving FIT uptake in 2007 and 2008 – CRC incidence has decreased and diagnosed disease has been down-staged. The FIT used is a single OC-SENSOR at 100 ng Hb/mL cut-off (equivalent to 20 µg Hb/g faeces); adherence of FIT+ve participants to colonoscopy is greater than 80%.

Jensen *et al* published a paper describing KP experience in California amongst 323,349 members who completed a FIT kit when invited for screening in 2007 or 2008 when aged 50-70 years (uptake 323,349/670,841 = 48.2%) and who were followed for up to four more annual screening episodes (7). Adherence to annual follow-up screening was high amongst those screened previously (between 75% and 86% at subsequent invitations) and median follow-up was four years. The proportion of FIT-screened participants with CRC who had positive FIT results in the year before the cancer was diagnosed (1,192 participants with CRC diagnosed within one year of FIT screening) was highest in the first round (84.5%), lower but stable in subsequent rounds (73.4% to 78.0%) and 80.4% overall after four rounds. FIT positivity (5.0%) was highest in round 1 and lower in subsequent rounds (3.7% to 4.3%). The PPVs for adenoma were highest in round 1 (51.5%) and were lower but stable in subsequent rounds (47.4% to 48.5%). The PPVs for CRC were also highest in the first round (3.4%) and lower but stable in subsequent rounds (2.1% to 2.3%). The authors concluded that annual FIT was associated with a high sensitivity for CRC, adherence to annual follow-up screening among initial participants was high and that annual programmatic FIT screening is feasible and effective for population-level CRC screening.

Data from the CISNET (Cancer Intervention and Surveillance Modeling Network), in conjunction with USPSTF, describe three independently-developed microsimulation models of CRC that are funded by the National Cancer Institute's CISNET consortium – *Simulation Model of Colorectal Cancer (SimCRC), Microsimulation Screening Analysis (MISCAN) for Colorectal Cancer, and Colorectal Cancer Simulated Population model for Incidence and Natural history (CRC-SPIN)*. Charts were presented illustrating colonoscopies and life-years gained (compared with no screening) for a cohort of 40-year-olds for FIT screening strategies that vary by age to begin, age to end, and screening interval, by model (39). For example, annual screening with FIT (green markers) starting at age 50 and raising the age to end screening from 75 to 80 years increased life-years gained by 5-7 per 1,000 (2-3%) while increasing the number of colonoscopies by 98-119 per 1,000 (6-7%) and the number of FITs by 1,618-1,709 per 1,000 (10-11%). Raising the age to end screening further, from age 80 to age 85, yielded even smaller gains in life-years (2-3 per 1,000, a 1% increase) relative to the change in the number of colonoscopies required (66-79 per 1,000, a 4% increase). The number of FITs increased by 1,162-1,244 per 1,000 (a 7% increase) (detailed observations reported by Zauber *et al* (39)).

Dutch investigators (40) compared the participation and diagnostic yield of repeated FIT testing with screening intervals of 1, 2 or 3 years. Positivity and detection of advanced neoplasia were significantly lower in the second screening round than in the first round of screening. There was no association between the interval length and the detection of advanced neoplasia at the second screening round. Second round participation was stable and acceptably high with screening intervals of 1-3 years.

Another study from the Netherlands (41) reported MISCAN-colon microsimulation modelling to estimate costs and benefits of FIT screening strategies with either 1- or 2-sample FIT screening, various FIT cut-offs for positivity and screening schedules that varied with respect to age range and screening interval. The authors concluded that increasing the screening intensity of 1-sample FIT testing (*i.e.* greater age range and/or shorter screening interval) is more cost-effective than providing two FIT within one screening round.

In conclusion, issues that drive differences between the studies include cut-off, colonoscopy resource and the definition of one year. In the KP California setting, it is reasonable to offer FIT screening every year, whilst in the Netherlands, biennial screening is appropriate.

**References**

1. Cole S, Smith A, Upton J, Young GP. Is single stool sampling as effective as two stool sampling in faecal immunochemical test based screening for colorectal cancer? Gastroenterology 2007;132(4 (Suppl. 2)):A-314.
2. van Roon AH, Wilschut JA, Hol L, van Ballegooijen M, Reijerink JC, t Mannetje H, et al. Diagnostic yield improves with collection of 2 samples in fecal immunochemical test screening without affecting attendance. Clin Gastroenterol Hepatol 2011;9(4):333-9.
3. Raginel T, Puvinel J, Ferrand O, Bouvier V, Levillain R, Ruiz A, et al. A population-based comparison of immunochemical fecal occult blood tests for colorectal cancer screening. Gastroenterology 2013;144(5):918-25.
4. Bowel Screening Pilot results: Ministry of Health – Manatū Hauora; 2016 [13/06/2016]. Available from: http://www.health.govt.nz/our-work/diseases-and-conditions/cancer-programme/bowel-cancer-programme/bowel-screening-pilot/bowel-screening-pilot-results.
5. van Ballegooijen M, Habbema JDF, Boer R, Zauber AG, Brown ML. AHRQ Technology Assessments. A comparison of the cost-effectiveness of fecal occult blood tests with different test characteristics in the context of annual screening in the Medicare population. Rockville (MD): Agency for Healthcare Research and Quality (US); 2003.
6. Imperiale TF, Ransohoff DF, Itzkowitz SH. Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med 2014;371(2):187-8.
7. Jensen CD, Corley DA, Quinn VP, Doubeni CA, Zauber AG, Lee JK, et al. Fecal immunochemical test program performance over 4 rounds of annual screening. A retrospective cohort study. Ann Intern Med 2016;164:456-63.
8. van der Meulen MP, Lansdorp-Vogelaar I, van Heijningen E-MB, Kuipers EJ, van Ballegooijen M. Nonbleeding adenomas: Evidence of systematic false-negative fecal immunochemical test results and their implications for screening effectiveness–A modeling study. Cancer 2016;122(11):1680-8.
9. Kapidzic A, Grobbee EJ, Hol L, van Roon AH, van Vuuren AJ, Spijker W, et al. Attendance and yield over three rounds of population-based fecal immunochemical test screening. Am J Gastroenterol 2014;109(8):1257-64.
10. Carroll MRR, Piggott C, Pearson S, Seaman HE, Halloran SP. Evaluation of quantitative faecal immunochemical tests for haemoglobin. Guildford, UK: Guildford Medical Device Evaluation Centre (GMEC), 2013. Available from: http://www.worldendo.org/assets/downloads/pdf/activities/fit_reports/gmec_fit_evaluation_report.pdf

11. Guittet L, Guillaume E, Levillain R, Beley P, Tichet J, Lantieri O, et al. Analytical comparison of three quantitative immunochemical fecal occult blood tests for colorectal cancer screening. Cancer Epidemiol Biomarkers Prev 2011;20(7):1492-501.
12. Pearson S, Bennitt W, Halloran S. Evaluation of eleven faecal occult blood test kits. London: Medical Devices Agency, 2000 MDA/2000/05
13. Daly JM, Bay CP, Levy BT. Evaluation of fecal immunochemical tests for colorectal cancer screening. J Prim Care Community Health 2013;4(4):245-50.
14. Wilschut JA, Hol L, Dekker E, Jansen JB, Van Leerdam ME, Lansdorp-Vogelaar I, et al. Cost-effectiveness analysis of a quantitative immunochemical test for colorectal cancer screening. Gastroenterology 2011;141(5):1648-55 e1.
15. Guittet L, Bailly L, Bouvier V, Launoy G. Indirect comparison of two quantitative immunochemical faecal occult blood tests in a population with average colorectal cancer risk. J Med Screen 2011;18(2):76-81.
16. Lee JK, Liles EG, Bent S, Levin TR, Corley DA. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. Ann Int Med 2014;160(3):171-81.
17. Launois R, Le Moine J-G, Uzzan B, Fiestas Navarrete LI, Benamouzig R. Systematic review and bivariate/HSROC random-effect meta-analysis of immunochemical and guaiac-based fecal occult blood tests for colorectal cancer screening. Eur J Gastroenterol Hepatol 2014;26(9):978-89.
18. Hundt S, Haug U, Brenner H. Comparative evaluation of immunochemical fecal occult blood tests for colorectal adenoma detection. Ann Intern Med 2009;150(3):162-9.
19. Faivre J, Dancourt V, Denis B, Dorval E, Piette C, Perrin P, et al. Comparison between a guaiac and three immunochemical faecal occult blood tests in screening for colorectal cancer. Eur J Cancer 2012;48(16):2969-76.
20. Brenner H, Tao S. Superior diagnostic performance of faecal immunochemical tests for haemoglobin in a head-to-head comparison with guaiac based faecal occult blood test among 2235 participants of screening colonoscopy. Eur J Cancer 2013;49(14):3049-54.
21. Tao S, Seiler CM, Ronellenfitsch U, Brenner H. Comparative evaluation of nine faecal immunochemical tests for the detection of colorectal cancer. Acta Oncol. 2013;52(8):1667-75.
22. Zubero MB, Arana-Arri E, Pijoan JI, Portillo I, Idigoras I, Lopez-Urrutia A, et al. Population-based colorectal cancer screening: comparison of two fecal occult blood test [sic]. Front Pharmacol 2014;4(175).
23. Chiang T-H, Chuang S-L, Li-Sheng Chen S, Chiu H-M, Ming-Fang Yen A, Yueh-Hsia Chiu S, et al. Difference in performance of fecal immunochemical tests with the same hemoglobin cut-off concentration in a nationwide colorectal cancer screening program. Gastroenterology 2014;147(6):1317-26.
24. Brignardello-Petersen R, Rochwerg B, Guyatt GH. What is a network meta-analysis and how can we use it to inform clinical practice? Pol Arch Med Wewn 2014;124(12):659-60.
25. Winawer SJ, Zauber AG, Ho MN, O'Brien MJ, Gottlieb LS, Sternberg SS, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med 1993;329(27):1977-81.
26. Obrien MJ, Winawer SJ, Zauber AG, Gottlieb LS, Sternberg SS, Diaz B, et al. The National Polyp Study - patient and polyp characteristics associated with high-grade dysplasia in colorectal adenomas. Gastroenterology 1990;98(2):371-9.
27. Winawer SJ, Zauber AG, Obrien MJ, Gottlieb LS, Sternberg SS, Stewart ET, et al. The National Polyp Study - design, methods, and characteristics of patients with newly diagnosed polyps. Cancer 1992;70(5):1236-45.
28. Winawer SJ, Zauber AG, Obrien MJ, Ho MN, Gottlieb L, Sternberg SS, et al. Randomized comparison of surveillance intervals after colonoscopic removal of newly diagnosed adenomatous polyps. N Engl J Med 1993;328(13):901-6.

29. Martínez ME, Baron JA, Lieberman DA, Schatzkin A, Lanza E, Winawer SJ, et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. Gastroenterology 2009 136(3):832-41.
30. Zauber AG, Winawer SJ, O'Brien MJ, Lansdorp-Vogelaar I, van Ballegooijen M, Hankey BF, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. New Engl J Med 2012;366(8):687-96.
31. Lane JM, Chow E, Young GP, Good N, Smith A, Bull J, et al. Interval fecal immunochemical testing in a colonoscopic surveillance program speeds detection of colorectal neoplasia. Gastroenterology 2010;139(6):1918-26.
32. Network NCICIaSM. Screening for colorectal cancer [01/06/2016]. Available from: http://cisnet.cancer.gov/projections/colorectal/screening.php.
33. Grazzini G, Ventura L, Zappa M, Ciatto S, Confortini M, Rapi S, 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(11):1511-5.
34. Cha JM, Lee JI, Joo KR, Shin HP, Park JJ, Jeun JW, et al. Performance of the fecal immunochemical test is not decreased by high ambient temperature in the rapid return system. Dig Dis Sci 2012;57(8):2178-83.
35. van Roon AHC, Hol L, van Vuuren AJ, Francke J, Ouwendijk M, Heijens A, et al. Are fecal immunochemical test characteristics influenced by sample return time? A population-based colorectal cancer screening trial. Am J Gastroenterol 2012;107(1):99-107.
36. Chausserie S, Levillain R, Puvinel J, Ferrand O, Ruiz A, Raginel T, et al. Seasonal variations do not affect the superiority of fecal immunochemical tests over guaiac tests for colorectal cancer screening. Int J Cancer 2015;136(8):1827-34.
37. van Rossum LG, van Rijn AF, van Oijen MG, Fockens P, Laheij RJ, Verbeek AL, et al. False negative fecal occult blood tests due to delayed sample return in colorectal cancer screening. Int J Cancer 2009;125(4):746-50.
38. Dancourt V, Hamza S, Manfredi S, Drouillard A, Bidan JM, Faivre J, et al. Influence of sample return time and ambient temperature on the performance of an immunochemical faecal occult blood test with a new buffer for colorectal cancer screening. Eur J Cancer Prev 2015;25(2):109-14.
39. Zauber A, Knudsen A, Rutter CM, Lansdorp-Vogelaar I, Kuntz KM, (Writing Committee of the Cancer Intervention and Surveillance Modeling Network (CISNET) Colorectal Cancer Working Group). Evaluating the benefits and harms of colorectal cancer screening strategies: a collaborative modeling approach. 2015 Contract No.: AHRQ Publication No. 14-05203-EF-2
40. van Roon AHC, Goede SL, van Ballegooijen M, van Vuuren AJ, Looman CWN, Biermann K, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut 2013;62(3):409-15.
41. Goede SL, van Roon AHC, Reijerink JCIY, van Vuuren AJ, Lansdorp-Vogelaar I, Habbema JDF, et al. Cost-effectiveness of one versus two sample faecal immunochemical testing for colorectal cancer screening. Gut 2013;62(5):727-34.