Washington 2015

7th Meeting of the Expert Working Group (EWG) – ‘FIT for Screening’

Friday, 15 May 2015: 10:00–12:00

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:

**Item 1.** Is Cologuard the new high bar for non-invasive CRC screening?  
*Chair: Robert E. Schoen, Professor of Medicine & Epidemiology, University of Pittsburgh, PA, USA*

Slide set no. 1  Cologuard: Yes. Robert E. Schoen
Slide set no. 2  Cologuard: No. Douglas J. Robertson, Geisel School of Medicine at Dartmouth, NH and VA Medical Center, VT, USA

**Item 2.** 80% national CRC screening rate by 2018: a goal we can reach by getting FIT  
*Chair: James E. Allison, Division of Gastroenterology, University of California San Francisco, USA*

Slide set no. 3  FIT and the FDA – the problems and possible solutions – a collegial discussion between stakeholders

Slide set no. 4  Industry Issues  
*FIT’s Quantity, Quality and Oversight*  
Helen Landicho, Senior Vice President Regulatory Affairs, Polymedco, Inc., USA

Slide set no. 5  Patient safety and quality laboratory issues  
*The Iowa experience with CLIA-waived FIT*  
Barcey Levy, Department of Family Medicine, University of Iowa, USA

Slide set no. 6  Possible FDA solutions for issues raised  
*FDA regulation of faecal immunochemical testing (FIT)*  
Yvonne Doswell, Scientific Reviewer, Center for Devices and Radiological Health, FDA, USA

**Item 3.** Free papers

Appendix 1  *The FIT Pilot in England*  
*Sally Benton, Director, NHS Bowel Cancer Screening Southern Programme Hub, Guildford, UK*

Slide set no. 7  Interval cancer after negative FIT and colonoscopy after positive FIT  
Isabel Portillo, Manager of Colorectal Detection Programme, Bilbao, Spain

Slide set no. 8  Determining an appropriate cut-off level for a national screening programme  
Ernst Kuipers, Professor of Medicine, Erasmus MC, Rotterdam, NL

Slide set no. 9  Colorectal cancer (CRC) screening in Ibaraki prefecture, Japan. A comparison between males and females using a two-day sampling method  
Yoko Saito, Ibarakiken Medical Center, Mito, Japan

Comparison of OC-SENSOR and FOB Gold in population-based colorectal cancer screening  
*Manon van der Vlugt, Department of Gastroenterology and Hepatology, Academic Medical Centre, Amsterdam, The Netherlands*

*Sally Benton was not able to attend the meeting. An abstract detailing outcomes from the FIT Pilot in England was included in the WEO CRC SC ‘Additional Material’ booklet; the abstract is reproduced here as Appendix 1.*
1. Is Cologuard the new high bar for non-invasive CRC screening?

Cologuard (Exact Sciences Corporation, Madison, WI 53719) is marketed as “the first and only FDA-approved stool DNA non-invasive colorectal cancer screening test”.¹ A single stool sample is collected into a container and then sent to Exact Sciences Labs for analysis (which includes quantitative molecular assays for KRAS mutations, aberrant NDRG4 and BMP3 methylation, and β-actin, plus a haemoglobin immunoassay [i.e. a faecal immunochemical test, FIT]). In 2014, Imperiale et al published the results of a study in the New England Journal of Medicine that compared a ‘Multitarget stool DNA test’ (Cologuard) with a single FIT (OC FIT-CHEK, Polymedco, USA; cut-off for positivity 20 µg haemoglobin/g faeces [100 ng Hb/mL buffer]) in a cohort of people at average risk for CRC who subsequently underwent colonoscopy.² Imperiale et al reported that:

*Of the 9989 participants who could be evaluated, 65 (0.7%) had colorectal cancer and 757 (7.6%) had advanced precancerous lesions (advanced adenomas or sessile serrated polyps measuring ≥1 cm in the greatest dimension) on colonoscopy. The sensitivity for detecting colorectal cancer was 92.3% with DNA testing and 73.8% with FIT (P=0.002). The sensitivity for detecting advanced precancerous lesions was 42.4% with DNA testing and 23.8% with FIT (P<0.001). The rate of detection of polyps with high-grade dysplasia was 69.2% with DNA testing and 46.2% with FIT (P=0.004); the rates of detection of serrated sessile polyps measuring 1 cm or more were 42.4% and 5.1%, respectively (P<0.001). Specificities with DNA testing and FIT were 86.6% and 94.9%, respectively, among participants with non-advanced or negative findings (P<0.001) and 89.8% and 96.4%, respectively, among those with negative results on colonoscopy (P<0.001). The numbers of persons who would need to be screened to detect one cancer were 154 with colonoscopy, 166 with DNA testing and 208 with FIT.*

*In asymptomatic persons at average risk for colorectal cancer, multitarget stool DNA testing detected significantly more cancers than did FIT but had more false positive results. (Funded by Exact Sciences; ClinicalTrials.gov number, NCT01397747.)*

*Slide set no. 1: Robert E. Schoen.

‘Cologuard: Pro’*

Professor Schoen presented an argument supporting the use of Cologuard, referring to his co-authored 2015 publication in Digestive Diseases and Sciences ‘Detection of Advanced Neoplasia with FIT versus sigmoidoscopy versus colonoscopy: more is more’.³ That publication concluded that colonoscopy detects more advanced neoplasia than FIT and/or flexible sigmoidoscopy (FS), although acknowledged that the number needed to screen makes colonoscopy expensive and inconvenient. The authors reported that FIT is least effective, adds nothing to a regime of FS and FIT and, when the focus is purely on detection, the data are clear: less is not more – more is more.

Professor Schoen also referred to work from Castro et al⁴ that concluded the diagnostic yield for advanced right-sided neoplasia is low with FIT and FS: “Fecal immunochemical test is more specific than sigmoidoscopy [for right-sided advanced colorectal neoplasia] but less sensitive than sigmoidoscopy [for the same] according to NORCCAP criteria [one distal polyp ≥ 10 mm, any adenoma, or CRC].”

Professor Schoen continued by outlining the results of the Imperiale et al study comparing Cologuard with FIT ² and emphasised the significantly greater detection of high grade dysplasia and sessile serrated adenoma with Cologuard.
The long-term effectiveness of CRC screening using a faecal occult blood test (FOBT) is realised by repeat testing; in Italy and the Netherlands with FIT,\textsuperscript{5,6} and in the UK with a guaiac-based FOBT,\textsuperscript{7} compliance may be less than ideal.

Professor Schoen concluded his presentation to the EWG by stating that sensitivity should be maximized whenever possible and that Cologuard is more sensitive than FIT for advanced adenomas and cancers.

**Slide set no. 2: Douglas J. Robertson.**

‘Is Cologuard the new high bar for non-invasive CRC screening? “No”’

Professor Robertson presented the argument that CRC screening using Cologuard may not be the best option. Important progress has been made by Thomas Imperiale and his co-workers since they reported in 2004 on DNA faecal testing that detected only 50% of invasive cancers and less than 20% of advanced adenoma.\textsuperscript{8} In 2014, Imperiale et al\textsuperscript{2} reported the results of further work comparing the performance of an updated panel of biomarkers and a FIT versus FIT alone. Sensitivities for the detection of CRC (93.2%) and advanced adenoma (42.4%) were much improved compared with the 2004 report\textsuperscript{8} and were significantly higher than for the one-time use of FIT alone (73.8% and 23.8%, respectively).\textsuperscript{9} Both FIT and Cologuard detected the Stage III/IV cancers (Cologuard detected a few more early stage cancers).

Importantly, Cologuard had a low specificity – about 10% of the cohort had a positive Cologuard result but an entirely negative colonoscopy, which is likely to prompt some clinicians to embark on further investigations.\textsuperscript{9}

Is Cologuard really likely to improve CRC mortality, when FIT repeated every one or two years will detect significant disease?

The practical considerations for use of Cologuard include compliance, follow-up and cost.

- FIT demonstrates a higher uptake than colonoscopy and, used just once, has been shown to detect the same number of cancers as colonoscopy (although more adenomas may be identified in the colonoscopy group).\textsuperscript{10}
- An entire stool is required for Cologuard analysis, whilst only a tiny sample is required for FIT. (It is important to note exclusions from the Imperiale study.\textsuperscript{2} Only 34 FIT samples were excluded because of insufficient sample compared with 689 stools samples taken for the DNA analysis.)
- The one-time cost of FIT is $20, which, if repeated annually would amount to $200 over 10 years. Cologuard costs about $500 for each test; three during the course of 10 years would cost $1500.

In summary, Cologuard is a more difficult test to complete and will have an impact on compliance; it detects too many false-positives (detecting non-advanced adenoma is a disadvantage) and cost is prohibitive. Only a long-term RCT comparing stool DNA testing with FIT, each at regular clinically defined intervals, would yield data that would be directly comparable.
2. ‘80% National CRC Screening Rate by 2018. A goal we can reach by getting FIT!’

*Slide set no. 3: James E. Allison*

Until 2012, colonoscopy every 10 years was promoted as the best strategy for CRC screening, FS was disparaged as clinically illogical, faecal tests were thought to limit opportunities for CRC prevention and compliance with repeat testing was considered poor. Since 2012, no screening test for CRC has been demonstrated to be superior to any other, although the preference for colonoscopy in the US amongst primary care providers is still evident. In November 2013 the Center for Disease Control’s Morbidity and Mortality Weekly Report (MMWR) promoted new key messages that encouraged primary care providers to offer all screening tests and promote the one that is most likely to be done, a message supported by the National Colorectal Cancer Roundtable (NCCRT), JAMA and by expert review.

Quantitative FIT have not yet been approved by the FDA for use in the USA. Only qualitative FIT are available and FIT approved for use in the USA are CLIA-waived (i.e. defined as a simple laboratory procedure that has an insignificant risk of an erroneous result). As of September 2014, there were 116 FDA-approved CLIA-waived FOBT for sale in the USA, very few with published evidence for performance and quality control. Professor Allison identified one qualitative FIT product available through Amazon.com for $22.99 which “is the same test used by physicians to effectively diagnose diverticulitis, colitis, ulcerative colitis and irritable bowel syndrome”!

FIT and the FDA – the problems and possible solutions – a collegial discussion between stakeholders

*Slide set no. 4: Helen Landicho (industry)*

‘FIT’s Quantity, Quality and Oversight’

According to the FDA, “Manufacturers must establish and follow quality systems to help ensure that their products consistently meet applicable requirements and specifications” and guidance for manufacturers since 2014 has improved. In 1988, the FDA passed CLIA (Clinical Laboratory Improvement Amendments) regulations to regulate laboratory testing and require clinical laboratories to be certificated before they can accept human samples for diagnostic testing. Laboratories can obtain multiple types of CLIA certificates, based on the kinds of diagnostic tests they conduct. CLIA regulations describe three levels of test complexity: waived tests, moderate complexity tests and high complexity tests.

According to the FDA database, there are 24 manufacturers of FIT and 37 differently labelled CLIA-waived FIT products. Fourteen manufacturers market multiple FIT and two manufacturers are linked with 13 different CLIA-waived FIT products. All rapid (qualitative) FIT are CLIA-waived.

Recommendations for the future: adopt CLSI (Clinical & Laboratory Standards Institutes) standards for FIT analytical performance measures, update FDA guidance and consider a certification programme for FIT (similar to that for NGSP [‘Harmonizing Hemoglobin A1C testing’]).
Slide set no. 5: Barcey Levy (physician)

‘The Iowa experience with CLIA-waived FIT’

Dr Levy outlined the results of a study that compared a single-sample CLIA-waived qualitative FIT with colonoscopy to report the sensitivity for proximal versus distal adenomatous polyps or cancer. All individuals scheduled for colonoscopy were invited to complete a FIT before the procedure. Exclusion criteria included familial polyposis syndromes, ulcerative colitis, Crohn’s disease or active rectal bleeding. Subjects without symptoms were included in the screening group. Subjects with previous polyps or colorectal cancer were included in the surveillance group. Individuals with a change in bowel habits, anaemia, positive FOBT or FIT, appetite change or abdominal pain were included in the diagnostic group. Invitations were issued to 2,336 and 1,026 ultimately completed a colonoscopy and FIT.

Over the course of the study, four different qualitative FIT products were used because of a change in name, then a product recall and then an unexpected decline in positivity (Polymedco’s OC-Light). The sensitivity of a single-sample qualitative FIT for advanced adenoma and adenocarcinoma was low (only 18%), although specificity was 90%. It is likely that primary care physicians are unaware of the limitations of CLIA-waived qualitative FIT.

Slide set no. 6: Yvonne Doswell (FDA)

‘FDA Regulation of Fecal Immunochemical Testing (FIT)’

Ms Doswell provided an overview of the FDA’s regulation of medical devices and challenges to regulating FIT. The FDA reviews submissions based on intended use and indications for use and classification of devices and evidence thereafter required for performance depends on risk (Class 1, low risk; Class II, moderate risk; Class III, high risk). Section 510(k) of the Food, Drug and Cosmetic Act requires device manufacturers who must register, to notify FDA of their intent to market a medical device at least 90 days in advance. This is known as Premarket Notification - also called PMN or 510(k). This allows FDA to determine whether the device is equivalent to a device already placed into one of the three classification categories. FIT are cleared through the 510(k) pathway and classified as Class II devices. A 510(k) submission requires analytical studies to determine performance characteristics, method comparisons using clinical samples and labelling with clear and concise instructions for use and interpretation of results. For a quantitative claim for FIT, the FDA encourages submission of study proposals and questions via the Pre-Submission programme. Pre-submission guidance is available here: http://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm311176.pdf
When FDA review is needed prior to marketing a medical device, FDA will either:

* "clear" the device (Class II) after reviewing a premarket notification, otherwise known as a 510(k) (named for a section in the Food, Drug, and Cosmetic Act), that has been filed with FDA, or

* "approve" the device (Class III; Class III devices are high risk devices that pose a significant risk of illness or injury, or devices found not substantially equivalent to Class I and II through the 510(k) process) after reviewing a premarket approval (PMA) application that has been submitted to the FDA.

Whether a 510(k) or a PMA application needs to be filed depends on the classification of the medical device.

To acquire clearance to market a device using the 510(k) pathway, the submitter of the 510(k) must show that the medical device is "substantially equivalent" to a device that is already legally marketed for the same use.

To acquire approval of a device through a PMA application, the PMA applicant must provide reasonable assurance of the device’s safety and effectiveness.

Slide set no. 7: Isabel Portillo

‘Interval cancer after negative FIT and colonoscopy after positive FIT’

The Basque Country provides population-based screening for men and women aged 50-69 years every two years using the OC-SENSOR device (Eiken Chemical Co. Ltd., Japan) with a cut-off of 20 µg Hb/g faeces (100 ng Hb/mL buffer). By 2014, the programme had achieved 100% coverage of the population. The investigators identified interval cancers diagnosed within two years after a negative FIT or after a positive FIT but negative colonoscopy. 77% of interval cancers were found after a negative FIT (all < 10 µg/g). Interval cancers were more common amongst people aged > 60 years, men and were more likely to be in the distal colon and rectum. Most interval cancers were at an advanced stage at diagnosis and were diagnosed within the first year after screening.

Slide set no. 8: Ernst Kuipers

‘Determining an appropriate cut-off level for a national screening programme’

Twenty four out of 28 EU countries now offer CRC screening (eight opportunistic, three piloting or planning a programme). Fifteen of those 24 countries offer FIT, but with a wide range in cut-offs for positivity (e.g. Belgium, Spain and New Zealand use OC-SENSOR with a cut-off of 15 µg Hb/g faeces), whilst Nova Scotia use a qualitative FIT (Hemoccult ICT, Beckman Coulter Inc., USA) that, according to the manufacturer, has an analytical sensitivity of 95% at 300 µg/g (online source, BR-17882A Hemoccult ICT Testing Highly Specific - Beckman Coulter [pdf]). As the cut-off for positivity increases, the positivity rate in the first round of screening decreases and the ‘number needed to screen’ and the ‘number needed to scope’ to detect one subject with advanced neoplasia decreases.
The cut-off selected by a screening programme does not influence uptake or the costs of the primary screening. However, by increasing the cut-off, the demand on colonoscopy resource will be less and the test will have a higher positive predictive value (lower number needed to scope), although some advanced neoplasia will be missed and a shorter interval between screening tests may be necessary. A decision must be made about an acceptable FIT positivity rate during initial rollout of a programme, and this can be derived using a simple equation:

\[
\text{Acceptable FIT positivity rate} = \frac{\text{colonoscopy capacity}}{\text{target population} \times \text{screening uptake} / \text{screening interval}}
\]

For example, with a colonoscopy capacity of 2,000 per year, a target population of 100,000, anticipated uptake of 60% and biennial screening:

\[
\text{Acceptable FIT positivity rate} = \frac{2,000}{(100,000 \times 0.6) / 2} = 6.6%.
\]

The desired positivity rate will determine the FIT cut-off used. The positivity rate and yield will decline over successive rounds of screening, which will influence decisions about FIT cut-off thresholds, screening interval and target age range. The FIT pilot in the Netherlands illustrates the issues.

Using OC-SENSOR at a cut-off of 10 µg Hb/g faeces (50 ng Hb/mL buffer), the Dutch FIT pilot had a FIT positivity of 6.4% and uptake of 50-62% amongst 50-75 year-olds. The number needed to scope for one advanced neoplasia was 1.8. In the first year of the Dutch national programme, using FOB Gold at a cut-off of 15 µg/g and inviting 63, 65, 67, 75 and 76 year-olds, uptake was 68%, FIT positivity 12% and number needed to scope 2.5. The challenge to limited colonoscopy resource was reduced in the second round of screening by increasing the cut-off to 47 µg/g, which yielded a more manageable positivity rate of 7.2% and number needed to scope of 2.1.

Data from modelling studies and pilots suggest that the most efficient approach during the steady state phase of a FIT screening programme is to use a low cut-off and adjust the screening interval\(^{17,19}\). Further research is needed to establish whether adjusting the target age range or using a two-sample FIT strategy at low cut-off might be useful.

**Slide set no. 9: Yoko Saito**

‘Colorectal cancer (CRC) screening in Ibaraki Prefecture, Japan. The comparison between males and females using a two-day sampling method’

CRC screening in Japan has been offered to men and women from the age of 40 years since 1992. The screening method varies by prefecture, although a two-day FIT method has been adopted widely. Uptake of screening is only about 35% (national target 40%). In the Ibaraki Prefecture, two-sample FIT screening is offered using the OC-SENSOR device (Eiken Chemical Co. Ltd., Japan) at a cut-off for positivity of 20 µg Hb/g faeces (100 ng Hb/mL buffer). Age-adjusted CRC incidence and mortality are lower for women than men. Because FIT positivity, the cancer detection rate and PPV are lower for women, should a lower FIT cut-off be used for women?
FIT performance data using a two-day or one-day sampling method during the period 2007-2012 were compared for men and women. For participants over the age of 50 years and using both sampling methods, FIT positivity was significantly greater for men than women (p=0.0000). There were statistically significant differences between the sexes in favour of men in cancer detection, PPV (for intra mucosal cancer) and detection of invasive cancers. However, the PPV for invasive cancer was not significantly different between men and women using either sampling method. The proportion of invasive cancers was the same for men and women and the investigators conclude that the FIT cut-off threshold should not be different for men and women.
Appendix 1

INCREASED PARTICIPATION IN COLORECTAL CANCER SCREENING DURING A PILOT OF A FAECAL IMMUNOCHEMICAL TEST FOR HAEMOGLOBIN (FIT) IN ENGLAND

S. Moss¹, C. Mathews¹, T. J. Day², S. Smith³, S. P. Halloran⁴,⁵

¹Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Queen Mary University of London, London, ²NHS Cancer Screening Programmes, Sheffield, ³NHS Bowel Cancer Screening Midlands and North West Programme Hub, Rugby, ⁴NHS Bowel Cancer Screening Southern Programme Hub, ⁵University of Surrey, Guildford, United Kingdom

Introduction: The NHS Bowel Cancer Screening Programme (BCSP) in England has used a guaiac faecal occult blood test (gFOBt) since 2006. In April 2014 the BCSP commenced a six-month FIT Pilot study to assess the clinical, financial and organisational implications of adopting FIT.

Aims & Methods: Two regional BCSP Hubs (Southern and Midlands & North West) and associated Screening Centres participated in the pilot study. One in 28 invitees was offered FIT rather than gFOBt. 30,000 FIT invitations provided adequate power for analysis of FIT uptake compared with gFOBt. The OC-SENSOR FIT system (Eiken Chemical Co. Ltd., Japan) was used with a cut-off for positivity of 20 µg haemoglobin [Hb]/g faeces (100 ng Hb/mL buffer).

Results: 40,930 subjects were invited to participate with a FIT and 1,126,087 with a gFOBt during the pilot period (April - October 2014). Uptake of FIT was significantly higher than gFOBt (66.5% vs. 59.4%; OR 1.36). The increase in uptake was significantly greater for previous non-responders (FIT 25.8% vs. gFOBt 14.2%; OR 2.09), compared with subjects invited for the first time (61.2% vs. 50.3%; OR 1.56) and those who had participated previously (90.3% vs. 86.1%; OR 1.50). The increase in uptake was higher in males (FIT 64.6% vs. gFOBt 56.4%; OR 1.41) than females (68.3% vs. 62.1%; OR 1.31) and was apparent for all quintiles of deprivation. Of particular note is the increase in uptake with FIT compared with gFOBt in the most deprived and traditionally ‘hard-to-reach’ quintile (53.7% vs. 45.8%; OR 1.37).

Overall positivity was 7.8% with FIT (cut-off 20 µg Hb/g faeces) and 1.7% with gFOBt (OR 4.83). The increase in positivity was similar in males and females and in all deprivation quintiles, but increased with age. Significantly more colorectal cancers (CRC) (0.27% FIT vs. 0.12% gFOBt; OR 2.19) and advanced adenomas (1.74% vs. 0.35%; OR 4.97) were detected with FIT. The PPV for all neoplasms was significantly higher with FIT (55.9% vs. 51.9%; OR 1.17). At a cut-off of 150 µg Hb/g faeces (750 ng Hb/mL buffer), which yielded a positivity for FIT (1.8%) similar to gFOBt, FIT had a higher detection rate and PPV for advanced adenomas and all neoplasms.

Conclusion: FIT significantly increased uptake of screening and provides an opportunity to adjust the faecal Hb concentration cut-off for positivity and thus the burden on colonoscopy resource. Further analysis will determine how the faecal Hb concentration measured by FIT could be incorporated into a multivariate risk score for CRC.
References