NICE advice

ColonFlag for identifying people at risk of colorectal cancer

Medtech innovation briefing
Published: 12 April 2018

www.nice.org.uk/guidance/mib142

Summary

- The **technology** described in this briefing is ColonFlag. It uses routinely available datasets to help identify people who are at high risk of developing colorectal cancer.
- The **innovative aspects** are that ColonFlag uses so-called big data and machine learning methods.
- The intended **place in therapy** would be as an addition to current investigations, such as faecal immunochemical tests, to help identify people who may need referral for suspected colorectal cancer.
- The **main points from the evidence** summarised in this briefing are from 4 observational studies including nearly 3.5 million patient records. They show that ColonFlag may provide an additional means for identifying people at risk of colorectal cancer, alongside standard screening and diagnostic tests.

- **Key uncertainties** around the evidence or technology are that there is no published evidence detailing the resource consequences of, or patient outcomes from, using ColonFlag.
- The **cost** of ColonFlag (excluding VAT) varies depending on the size of the population served per installation. For example, the annual estimated per capita cost in year 1 of an installation covering a population of 300,000 is £1.30. The **resource impact** would be an initial increased cost to the NHS including the time spent managing the system.

The technology

ColonFlag (Medial EarlySign) is a web-based machine learning algorithm that is designed to help identify people aged 40 years or over who are at high risk of having colorectal cancer (CRC).

The algorithm uses existing datasets containing age, sex and complete blood count (CBC) test results to generate a risk score for a person. CBCs are widely used and measure the amount of red and white blood cells, haemoglobin and other factors in the blood. People with a high-risk score can be referred for further assessment, potentially before they show any symptoms.

ColonFlag works best when multiple CBCs, taken at different times, are available. The algorithm uses decision trees (similar to flow charts) and gradient boosting (which increases the prediction power of decision trees) to assess trends in the CBC parameters and patient demographics.

ColonFlag software would be installed on the central computer system within a healthcare organisation or network. The installation may be on the laboratory information system of a central laboratory or at a central electronic medical records system, such as EMIS or SystmOne. It produces a test report containing a risk score with a confirmation message.

If the score is above a defined threshold, the software will show that a person is at high risk of having CRC. The information will be passed on to the person's GP or the organisation's cancer lead. Minimal training may be needed to interpret the risk score; this is provided by the company.

ColonFlag automatically processes new data once it is saved in the healthcare network database.

Innovations

ColonFlag differs from other risk-prediction methods for colorectal cancer by automatically testing people using routinely available data. It differs from other software algorithms by using machine learning techniques to improve the accuracy of predictions.

Current NHS pathway or current care pathway

The NHS bowel cancer screening programme offers a guaiac faecal occult blood test (gFOBT) to people aged between 60 and 74 once every 2 years. In 2015, the UK National Screening Committee recommended that the bowel cancer screening programme switch to using a faecal immunochemical test (FIT) for primary screening.

For people with symptoms, NICE's diagnostics guidance on quantitative faecal immunochemical tests for colorectal cancer in primary care recommends 3 FITs. These tests are recommended for guiding referral in people without rectal bleeding and who do not meet the criteria for a suspected cancer referral outlined in NICE's guideline on suspected cancer. Other investigative tests, such as a barium enema or a flexible sigmoidoscopy, may also be done. A positive investigative test should be followed by a biopsy for diagnostic proof, and staging is done using contrast-enhanced CT.

Treatment can then begin based on the stage of the cancer.

Population, setting and intended user

ColonFlag is intended to identify, in addition to other testing methods, people who have an above-average risk of having CRC. Each software installation would cover the population of the host healthcare organisation. An administrator would be needed to operate the central computer and they would pass on the records of people with high-risk scores to either the person's GP or the organisation's cancer lead. The GP or cancer lead would then refer people to a gastroenterologist for a colonoscopy, unless contraindicated by a person's history or other risk factors.

Costs

Technology costs

Assuming that 50% of the population are aged 40 years or over (Office for National Statistics 2017), 27 million people would be eligible for testing using ColonFlag. The cost per person will decrease if the population covered per installation is higher (see table 1).

Table 1 Cost of ColonFlag

Description	Year 1 cost (£; excluding VAT)	Year 2 onwards cost (£; excluding VAT)
Installation fee (1 time)	50,000	0
Technical and clinical support (annual) per installation	40,000	40,000
Maintenance cost*	0	0
Cost per person tested**	1 to 5	1 to 5
Annual cost per capita for a population of 30,000 (£5 per person tested)	8	6.33
Annual cost per capita for a population of 300,000 (£1 per person tested)	1.30	1.13
* Technology runs on a standard internet information services server and has no additional maintenance cost. ** The annual price per person tested is lower in larger populations.		

Costs of standard care

NICE's diagnostics guidance on quantitative faecal immunochemical tests to guide referral for colorectal cancer in primary care gives the cost of a gFOBT as £0.78 and the cost of a FIT to be between £1.96 and £6.04. A colonoscopy is listed as costing £372.

For people outside the screening programme's age range, the standard care would be

clinical judgement in a standard GP appointment (£36, Curtis and Burns 2016).

Resource consequences

Using ColonFlag would add initial installation and implementation costs and is anticipated to increase both the number of referrals to gastroenterology and colonoscopies performed. This would need more GP resources to make referrals as well as more gastroenterologist and administrator resources to deal with more referrals. However, if more cancers were identified at an earlier stage, using ColonFlag could improve outcomes and may reduce the resources needed for cancer treatment and support. No published economic evidence on the resource consequences of adopting the technology was identified.

Minimal information technology integration is needed and basic information technology training is provided, free of charge. Basic clinical training is provided to help to interpret the risk score.

ColonFlag is not currently used in any UK organisation.

Regulatory information

ColonFlag was CE marked as a class I device in March 2017.

Equality considerations

NICE is committed to promoting equality, eliminating unlawful discrimination and fostering good relations between people with particular protected characteristics and others. In producing guidance and advice, NICE aims to comply fully with all legal obligations to: promote race and disability equality and equality of opportunity between men and women, eliminate unlawful discrimination on grounds of race, disability, age, sex, gender reassignment, marriage and civil partnership, pregnancy and maternity (including women post-delivery), sexual orientation, and religion or belief (these are protected characteristics under the Equality Act 2010).

Three quarters of colorectal cancer (CRC) cases are in people aged 65 or over. Jewish people of central and eastern European family origin are also thought to be at increased

risk. Age, religious beliefs and ethnicity are protected characteristics under the Equality Act 2010. People with cancer are protected under the Equality Act from the point of diagnosis.

Clinical and technical evidence

A literature search was carried out for this briefing in accordance with the interim process and methods statement. This briefing includes the most relevant or best available published evidence relating to the clinical effectiveness of the technology. Further information about how the evidence for this briefing was selected is available on request by contacting mibs@nice.org.uk.

Published evidence

Four studies including 3,485,065 patient records are summarised in this briefing. Patient records were drawn from populations in Israel, the UK and the US. The most recent study (Birks et al. 2017), reports the results of a large UK study of 2,550,119 patient records. Table 2 summarises the clinical evidence as well as its strengths and limitations.

Overall assessment of the evidence

There are 2 studies performed on data from GPs in the UK, which account for 75% of the data in this briefing. Most of the evidence compares the diagnostic accuracy of ColonFlag to standard care, which may vary across country and region. The area under the receiver operator curve (AUC) and the odds ratio (OR) of having colorectal cancer are the most commonly reported outcomes. These showed that the overall performance of the algorithm was reasonably consistent across the different populations. The AUC was slightly lower in the UK study, because the time interval before diagnosis was longer than the other studies. In an age-matched, case-control design from the same study, the AUC was considerably lower than in the other studies, showing that age is an important predictive factor. The reported ORs showed that ColonFlag is potentially useful for identifying people at 10 to 30 times increased risk of colorectal cancer (CRC).

All of the studies are retrospective and observational. There is no evidence available on the cost effectiveness, resource consequences or utility of ColonFlag.

Table 2 Summary of selected studies

Kinar et al. (2016)	
Study size, design and location	A retrospective, observational study on registry data in Israel (n=779,654) and the UK (n=25,613). All people aged 40 or over within the Maccabi Healthcare Services who had CBC results from 2008 and 2009 were included. The UK dataset was comprised of a subset of an anonymised UK primary care database. This resulted in a cohort design for the Israeli dataset and a case-control design for the UK dataset; 80% of the Israeli cohort was used as a derivation dataset and the remaining 20% was used as a validation dataset.
Intervention and comparator(s)	ColonFlag compared with the standard of care and gFOBT.
Key outcomes	**Israeli dataset:** AUC=0.82±0.01. OR at a false-positive rate of 0.5% was 26±5 and the specificity at 50% sensitivity was 88±2%. **UK dataset:** AUC=0.81. OR at a false-positive rate of 0.5% was 40±6 and the specificity at 50% sensitivity was 94±1%. ColonFlag detected 48% more CRC cases than gFOBT in a dataset of 75,822 Israeli records.
Strengths and limitations	A large number of patient records were included across 2 populations; the Israeli cohort was randomised, the UK one was not. The reference standards used are not clear (and may vary over the dataset), other than the comparison between ColonFlag and gFOBT in the Israeli subset. No power calculation is reported.
Birks et al. (2017)	

Study size, design and location	A retrospective, observational study on registry data from the Clinical Practice Research Datalink (CPRD) in the UK (n=2,550,119). People aged 40 or over with a CBC result from January 2000 to April 2015 were included. Following the methodology used in Kinar (2016), a primary analysis (n=2,225,249) and sensitivity analyses were performed. The sensitivity analyses included a cohort study, performed for people with CBCs taken during 2012 (n=600,273) and a case-control study, matching for age, sex and year of risk score (n=519,241).
Intervention and comparator(s)	ColonFlag compared with the standard of care.
Key outcomes	AUC=0.776 (95% CI 0.771 to 0.781) for CBCs taken in an 18–24 month interval before diagnosis. For the case-control group (age-matched), the AUC was 0.583 (95% CI 0.574 to 0.591). In the 2012 cohort, the PPV was 8.8% and the NPV was 99.6% at a specificity of 99.5%. At this cut-off, the OR was found to be 26.5 (95% CI 23.3 to 30.2). The AUC was 0.781 (95% CI 0.772 to 0.791).
Strengths and limitations	A very large study population within the NHS was used, including a large cohort group – allowing PPV and NPV to be calculated. A sensitivity analysis was performed showing that the results were robust to variations in randomly selected CBCs. Some of the individuals with no diagnosis during the study period may have been diagnosed outside the follow-up interval. The reference standards may have varied among patients. No power calculation is reported.
Hornbrook et al. (2017)	
Study size, design and location	A retrospective, observational study on registry data in the US (n=17,095). Data from people with CRC and CBC results before diagnosis (from 2000 to 2013; n=900) were included, along with data from 16,195 controls. The data were taken from the KPNW tumour registry.
Intervention and comparator(s)	ColonFlag compared with the standard of care.

Key outcomes	AUC=0.80±0.01. OR=34.7 (95% CI 28.9 to 40.4) for a specificity of 99%. ColonFlag was found to be more accurate at detecting right-sided CRCs than left-sided tumours.
Strengths and limitations	A smaller population was included than in the other studies and all records were drawn from within a single private healthcare service. This study uses a smaller age range (40 to 89 years) than the ColonFlag intended use (>40). Reference standards varied between records and may have included multiple tests (colonoscopies, flexible/ rigid sigmoidoscopies, gFOBT and FIT). No power calculation is reported.
Kinar et al. (2017)	
Study size, design and location	A retrospective, observational study on registry based data in Israel (n=112,584) for all people aged between 50 and 75 years within the Maccabi Healthcare Services with CBC results from July to December 2007.
Intervention and comparator(s)	ColonFlag compared with the standard of care.
Key outcomes	ColonFlag risk scores were converted into percentiles. 3,337 individuals were within a 3-percentile cut-off and 1,094 within a 1-percentile cut-off. In the 3-percentile group, the OR was 10.9 (95% CI 7.3 to 16.2), sensitivity was 25%. In the 1-percentile group, the OR was 21.8 (95% CI 13.8 to 34.2), sensitivity was 17.3%. Anticoagulant treatments, treatments for other gastrointestinal diseases and presence of other cancers were found to be possible causes of false positives.
Strengths and limitations	A small population was used in comparison to the other studies. The age group used was different than in the other studies and the population was all drawn from the same private healthcare service. There was no information given on the reference standard. No power calculation is reported.

Abbreviations: AUC, area under the receiver operator curve; CBC, complete blood count; CRC, colorectal cancer; gFOBT, guaiac faecal occult blood test; FIT, faecal immunochemical test; CI, confidence interval; KPNW, Kaiser Permanente Northwest region; OR, odds ratio; PPV, positive predictive value; NPV, negative predictive value.

Table 3 Summary of findings

Author	Location	Number of patients	AUC	OR of colorectal cancer
Kinar et al. 2016	Israel and UK	805,267	**Israel**: 0.82±0.01 **UK**: 0.81	**Israel**: (at false-positive rate 0.5%) 26±5 **UK**: (at false-positive rate 0.5%) 40±6
Birks et al. 2017	UK	2,550,119	**Primary Analysis**: 0.776 (95% CI 0.771 to 0.781) **Case-control group (age-matched):** 0.583 (95% CI 0.574 to 0.591) **Cohort**: 0.781 (95% CI 0.772 to 0.791)	26.5 (95% CI 23.3 to 30.2; for a specificity of 99.5%)
Hornbrook et al. 2017	US	17,095	0.80±0.01	34.7 (95% CI 28.9 to 40.4; for a specificity of 99%)
Kinar et al. 2017	Israel	112,584	–	**3-percentile group**: 10.9 (95% CI 7.3 to 16.2) **1-percentile group**: 21.8 (95% CI, 13.8 to 34.2)
Abbreviations: AUC, area under the receiver operator curve; CI, confidence interval; OR, odds ratio.				

Recent and ongoing studies

- Prediction of findings at screening colonoscopy using a machine learning algorithm based on complete blood counts (ColonFlag): Robert J Hilsden, Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada. Status: submitted for publication.
- Computer-assisted flagging of individuals at high risk of colon cancer in a large health maintenance organization using the ColonFlag Test: R. Goshen – Medial EarlySign Varda Shalev – Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. Status: submitted for publication.
- Validation of the model's performance in the detection of colorectal cancer and precancerous findings, based on cancer registry and colonoscopy data, Kaiser Permanente North California, US. Status: submitted for publication.

Specialist commentator comments

Comments on this technology were invited from clinical experts working in the field and relevant patient organisations. The comments received are individual opinions and do not represent NICE's view.

Four of the 6 experts were familiar with risk assessment software for colorectal cancer but none had used ColonFlag before.

Level of innovation

Two experts agreed that the technology application is innovative but that the technology itself is not. Two experts thought that the machine learning component is innovative but another thought that more advanced machine learning techniques are available. Two experts believed that the concept is novel but unproven. Five out of 6 experts said there are no technologies that predicted risk in asymptomatic patients but 1 stated that there are. This commentator was not convinced that machine learning techniques are better at predicting risk than other methods.

Three experts were aware of technologies that are similar to ColonFlag, but 2 added that they may be better for diagnostic rather than prognostic purposes because the risk factors assessed are different. One expert highlighted that in primary care, risk-prediction

algorithms are already used for colon cancer and other long-term conditions to help patients and clinicians with decision-making.

Potential patient impact

Four experts thought that using ColonFlag may lead to earlier diagnosis and treatment for colorectal cancer, which would improve patient survival and quality of life. Two experts thought ColonFlag being non-invasive could lead to fewer anxious patients. One expert thought that ColonFlag would personalise patient care. However, 1 expert thought these benefits would only happen if every patient in a GP practice had a complete blood count (CBC) test each year. One expert thought that using ColonFlag may lead to more guaiac faecal occult blood test (gFOBT) or faecal immunochemical test (FIT) tests, and another thought that it would lead to more colonoscopies.

Three experts thought that people outside the age range for screening but who are at high risk of colorectal cancer would benefit most from ColonFlag. One expert believed that all at-risk groups for colorectal cancer would benefit, whereas another expert said that older people would benefit. One expert thought that all people over 40 would potentially benefit but particularly people without symptoms. Two of the experts thought that up to half of the population of England would be eligible for ColonFlag but that its use would depend on the proportion of people with CBC results.

Potential system impact

Three experts thought that ColonFlag had the potential to reduce costs for the NHS by avoiding unnecessary investigations, decreasing treatment costs and providing more efficient care. Another expert thought that ColonFlag could lead to more consistency in gFOBT/FIT testing in primary care.

Two experts thought that ColonFlag would represent an additional cost to the NHS. Three experts felt that ColonFlag was expensive with 1 expert stating that gFOBT/FIT testing was cheaper and 2 stating the maintenance charge of £40,000 was excessive. One expert thought that the cost would be around the same as current practice but with improved targeting of resources. One expert thought that the overall cost effect was unknown because more patients would need treatment but fewer patients would need later-stage treatments.

One expert thought that ColonFlag would move care and resources upstream by identifying cancer earlier. Another expert noted that more staff would be needed for consultations with patients at high risk. Further, more resources would be needed for the increased use of the next test in the pathway. One expert thought that the resource impact would depend on the new care pathway; they stated that if gFOBT/FIT testing was the next step on the care pathway, the effect would be a small extra cost. If colonoscopy was the next step, there would be a larger increase in cost. Another expert agreed that there would be considerable resource consequences in secondary care to provide the extra investigations. Another expert remarked that adopting ColonFlag would need resources in primary care, such as extra GP time or hiring an administrator to operate the system. One expert did not foresee any changes to resource use from adopting ColonFlag.

The experts largely agreed that there would be little need for change to infrastructure to use ColonFlag. Three experts thought that only minimal training would be needed. One expert thought that there may need to be changes to infrastructure if there was insufficient capacity to deal with increased demand for colonoscopies. Another expert thought that some changes to infrastructure may be needed to incorporate the technology with electronic health records.

Two experts expressed concerns about patient confidentiality and stressed the importance of keeping patient details secure. Another expert thought that the technology itself had not been evaluated thoroughly and independently and could miss patients at high risk. ColonFlag would be an addition to the current standard of care but 1 expert said that there is a lot of variation in how frequently GPs use risk assessment tools.

General comments

One expert was unsatisfied with ColonFlag's reporting, particularly that statistical assessment of the tool was impossible because it is a closed-source software. The expert added that ColonFlag's reporting does not follow the standard transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) guidelines and there was no comparison of the algorithm to other methods like logistic regression. Furthermore, the expert had concerns about the applicability of the algorithm to the UK population. They thought that recalibrating the model to UK data would improve outcomes.

Three experts raised concerns about the cost of ColonFlag as a barrier to its adoption, with one also considering time constraints to be an obstacle. Another expert thought that there was a need to discuss ColonFlag's place in the patient pathway. ColonFlag's

incorporation within the software systems currently used in the NHS was also considered to be a potential issue by 1 expert.

All of the experts felt that more research was needed to address the uncertainties in the evidence base.

Specialist commentators

The following clinicians contributed to this briefing:

- Matthew Sperrin, senior lecturer in health data science, University of Manchester. No conflicts of interest declared.
- Joseph Huang, consultant colorectal surgeon, Queen's Hospital, Romford. No conflicts of interest declared.
- Niels Peek, professor of health informatics, University of Manchester. No conflicts of interest declared.
- Roger Motson, colorectal surgeon and director of the ICENI Centre, Colchester Hospital. Adviser to Cambridge Medical Robotics. The ICENI Centre receives educational grants from many companies in the medical device industry.
- Vardhini Vijay, consultant surgeon, Princess Alexandra Hospital, Harlow. No conflicts of interest declared.
- Juliet Usher-Smith, clinical senior research associate and GP, Primary Care Unit, University of Cambridge. No conflicts of interest declared.

Development of this briefing

This briefing was developed for NICE by King's Technology Evaluation Centre (KiTEC). The interim process and methods statement sets out the process NICE uses to select topics, and how the briefings are developed, quality-assured and approved for publication.

ISBN: 978-1-4731-2927-6