The fecal hemoglobin concentration, age and sex test score

on behalf of the COLONPREDICT study investigators

Published in:
International Journal of Cancer

DOI:
10.1002/ijc.30639

Publication date:
2017

Document Version
Peer reviewed version

The FAST (Faecal Haemoglobin Concentration, Age and Sex Test) Score

Cubiella, Joaquín; Digby, Jayne; Rodríguez-Alonso, Lorena; Vega, Pablo; Salve, María; Díaz-Ondina, Marta; Strachan, Judith; Mowat, Craig; Mcdonald, Paula; Carey, Francis; Godber, Ian M.; Ben Younes, Hakim; Rodriguez-Moranta, Francisco; Quintero, Enrique; Álvarez-Sánchez, Victoria; Fernández-Bañares, Fernando; Boadas, Jaume; Campo, Rafel; Bujanda, Luis; Garayoa, Ana; Ferrandez, Ángel; Piñol, Virginia; Rodríguez-Alcalde, Daniel; Guardiola, Jordi; Steele, Robert; Fraser, Callum; on behalf of the COLONPREDICT study investigators

Published in:
International Journal of Cancer

DOI:
10.1002/ijc.30639

Publication date:
2017

Document Version
Peer reviewed version

THE FAST (FAECAL HAEMOGLOBIN CONCENTRATION, AGE AND SEX TEST) SCORE: DEVELOPMENT AND EXTERNAL VALIDATION OF A SIMPLE PREDICTION TOOL FOR COLORECTAL CANCER DETECTION IN SYMPTOMATIC PATIENTS

Joaquín Cubiella[1,2], Jayne Digby[3], Lorena Rodríguez-Alonso[4], Pablo Vega[1], María Salve[1], Marta Díaz-Ondina[5], Judith A Strachan[6], Craig Mowat[7], Paula J McDonald[8], Francis A Carey[9], Ian M Godber[10], Hakim Ben Younes[11], Francisco Rodriguez-Moranta[4], Enrique Quintero[12], Victoria Álvarez-Sánchez[13], Fernando Fernández-Bañares[14], Jaume Boadas[15], Rafel Campo[16], Luis Bujanda[17], Ana Garayoa[18], Ángel Ferrandez[19], Virginia Piñol[20], Daniel Rodríguez-Alcalde[21], Jordi Guardiola[4], Robert JC Steele[3] and Callum G Fraser[3] on behalf of the COLONPREDICT study investigators[22]

Running title: FAST Score prediction tool

Authors' affiliations

[1]Gastroenterology Department, Complexo Hospitalario Universitario de Ourense, Ourense. Spain.

[2]Instituto de Investigación Biomedica (IBI) Ourense, Pontevedra y Vigo, Vigo. Spain.

[3]Centre for Research into Cancer Prevention and Screening, University of Dundee, Ninewells Hospital & Medical School, Dundee, Scotland, UK

[4]Department of Gastroenterology and Hepatology, University Hospital of Bellvitge-IDIBELL, L'Hospitalet de Llobregat, Barcelona, Spain; Ciber de Epidemiología y Salud Pública (CIBERESP), Spain.

[5]Clinical Analysis Department, Complexo Hospitalario Universitario de Ourense. Ourense. Spain.

[6]Blood Sciences, Ninewells Hospital & Medical School, Dundee, Scotland, UK.

[7]Department of Gastroenterology, Ninewells Hospital & Medical School, Dundee, Scotland, UK.

[8]Scottish Bowel Screening Centre, Kings Cross Hospital, Dundee, Scotland, UK

[9]Department of Pathology, Ninewells Hospital & Medical School, Dundee, Scotland, UK.

[10]Department of Biochemistry, Monklands Hospital, Airdrie, Lanarkshire, Scotland, UK.

[11]Department of Surgery, Wishaw General Hospital, Wishaw, Lanarkshire, Scotland, UK.

[12]Gastroenterology Department, Hospital Universitario de Canarias, Instituto Universitario de Tecnologías Biomédicas (ITB) & Centro de Investigación Biomédica de Canarias (CIBICAN), Universidad de La Laguna. Tenerife, Spain.

[13]Gastroenterology Department, Complejo Hospitalario de Pontevedra, Pontevedra, Spain.

[14]Gastroenterology Department, Hospital Universitari Mútua de Terrassa, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd). Terrassa, Spain.

[15]Gastroenterology Department, Consorci Sanitari de Terrassa, Terrassa, Spain.

[16]Gastroenterology Department, Hospital de Sabadell, Corporació Sanitària i Universitària Parc Taulí, Sabadell, Spain

[17]Donostia Hospital, Biodonostia Institute, University of the Basque Country UPV/EHU. (CIBERehd). San Sebastian. Spain.

[18]Gastroenterology Department, Hospital de Sagunto, Sagunto, Valencia, Spain.

[19]Servicio de Aparato Digestivo, Hospital Clínico Universitario, IIS Aragón, University of Zaragoza, (CIBERehd). Zaragoza, Spain.

[20]Gastroenterology Department, Hospital Dr. Josep Trueta, Girona, Spain.

[21]Digestive Disease Section, Hospital Universitario de Móstoles, Madrid, Spain.

[22]All investigators are listed in the Supplementary note.

Address for correspondence: Dr. Joaquín Cubiella**.** Department of Gastroenterology. Complexo Hospitalario Universitario de Ourense. C/ Ramón Puga 52-54. 32005. Ourense, Spain; phone number: +34 988385399**;** fax number: +34 988385399**;** e-mail: joaquin.cubiella.fernandez@sergas.es

Word count

Abstract: 250

Text: 4,320

Tables: 5

Supplementary tables: 1

Figures: 1

Key words**:** advanced colorectal neoplasia, colonoscopy, colorectal cancer, diagnostic accuracy, faecal immunochemical test, inflammatory bowel disease, risk stratification.

Abbreviations:

AN, advanced neoplasia; AUC, area under the curve; CI, confidence interval; CRC, colorectal cancer; EPAGE, European Panel on the Appropriateness of Gastrointestinal Endoscopy; f-Hb, faecal haemoglobin concentration; FIT, faecal immunochemical test for haemoglobin, IBD; inflammatory bowel disease, LIMS; laboratory information management system, NICE, National Institute for Health and Care Excellence; NNS, number needed to scope; OR, odds ratio; PPV, positive predictive value; ROC, receiver operating characteristic; SCL, significant colonic lesion.

ABSTRACT

Prediction models for colorectal cancer (CRC) detection in symptomatic patients, based on easily obtainable variables such as faecal haemoglobin concentration (f-Hb), age and sex, may simplify CRC diagnosis. We developed, and then externally validated, a multivariable prediction model, the FAST Score, with data from five diagnostic test accuracy studies that evaluated quantitative faecal immunochemical tests in symptomatic patients referred for colonoscopy. The diagnostic accuracy of the Score in derivation and validation cohorts was compared statistically with the area under the curve (AUC) and the Chi-square test.1,572 and 3,976 patients were examined in these cohorts, respectively. For CRC, the odds ratio (OR) of the variables included in the Score were: age (years): 1.03 (95% confidence intervals (CI): 1.02-1.05), male sex: 1.6 (95% CI: 1.1-2.3) and f-Hb (0-<20 μg Hb/g faeces): 2.0 (95% CI: 0.7-5.5), (20-<200 μg Hb/g): 16.8 (95% CI: 6.6-42.0), ≥200 μg Hb/g: 65.7 (95% CI: 26.3-164.1). The AUC for CRC detection was 0.88 (95% CI: 0.85-0.90) in the derivation and 0.91 (95% CI: 0.90-093; p = 0.005) in the validation cohort. At the two Score thresholds with 90% (4.50) and 99% (2.12) sensitivity for CRC, the Score had equivalent sensitivity, although the specificity was higher in the validation cohort (p<0.001). Accordingly, the validation cohort was divided into three groups: high (21.4% of the cohort, positive predictive value - PPV: 21.7%), intermediate (59.8%, PPV: 0.9%) and low (18.8%, PPV: 0.0%) risk for CRC. The FAST Score is an easy to calculate prediction tool, highly accurate for CRC detection in symptomatic patients.

Novelty and Impact

Lower gastrointestinal symptoms are very common in patients presenting in primary care, but colorectal cancer (CRC) is much rarer. Faecal haemoglobin concentration (f-Hb), at low thresholds, is a good rule-in test for CRC, but is dependent on age and sex. The FAST Score, based on the easily available variables of f-Hb, age and sex, has high diagnostic accuracy for CRC detection in symptomatic patients. It can be applied to generate three easy to interpret categories of risk to simplify referrals for colonoscopy.

INTRODUCTION

Colorectal cancer (CRC) remains an important health problem in Western Europe. It is the seventh cause of death and the fourth cause of years of life lost.[1] Health authorities have developed two main strategies to reduce the impact of CRC: screening in average and, on occasion, high-risk (personal or family history of CRC or adenomas) populations, and prompt detection in symptomatic patients.[2–6]

Although screening programmes are being progressively implemented, most CRC are detected when symptoms become apparent.[7] In addition, although gastrointestinal symptoms are extremely common in the population, the probability of CRC detection associated with any one symptom is low.[8–10] Which symptomatic patients should be evaluated promptly, mainly with colonoscopy, is the dilemma in any diagnostic strategy. Over-investigation generates financial costs and patients are subjected to psychological impact and unnecessary endoscopy-related risk and discomfort. On the other hand, lack of investigation risks a delay in diagnosis with clinical and possibly medicolegal impact.

In order to reach a balance, several approaches that use criteria for referral for colonoscopy that are associated with a high risk for CRC have been established. In this regard, the best known guidelines are the very detailed and extremely prescriptive National Institute for Health and Care Excellence (NICE) criteria for suspected cancer.[3,11] Additionally, a number of CRC prediction models have been designed and validated in different settings. These prediction models, systematically reviewed in detail recently, are calculated from

mathematical equations based mainly on symptoms, although information regarding demographics, other variables and results of investigations are sometimes also included.[12] Although their diagnostic accuracy is generally deemed to be satisfactory and better than the existing referral criteria, these prediction models have not been widely implemented, in part due to their complexity and the difficulty of collection of all of the spectrum of variables required.[13–15]

Faecal immunochemical tests for haemoglobin (FIT) have proven to be the best currently available non-invasive test for CRC screening in asymptomatic individuals and an excellent test for rule-in of CRC and rule-out of significant colonic lesions (SCL) in patients presenting with lower gastrointestinal symptoms.[16–21] Quantitative FIT allow for assessment of faecal hemoglobin concentration (f-Hb). There are several prediction models in asymptomatic individuals for CRC screening based on f-Hb,[22] and a CRC prediction model in symptomatic patients including f-Hb has been recently developed and externally validated.[23] This prediction model was shown to have high diagnostic accuracy and discriminated between three risk groups, one of them with a negligible risk for CRC. However, its applicability may be limited due to its complexity, since the model includes eleven variables.

In consequence, we carried out a derivation and external validation multivariable prediction model study, based on the hypothesis that a simple to calculate risk score, based on only three readily available variables - f-Hb, age and sex - could simplify the evaluation of symptomatic patients and provide a easy to use tool for the risk stratification of these to facilitate prioritisation of referrals for endoscopy.[20] In order to perform this study on the FAST Score,

detailed data obtained in five previously peer-reviewed and published studies evaluating the diagnostic accuracy of FIT in 5,548 symptomatic patients referred for colonoscopy were used.[16–21,23]

PATIENTS AND METHODS

1. Design

The development and external validation of the multivariable prediction model was designed according to the TRIPOD guidelines.[24] All diagnostic accuracy studies from which data were derived for this work were performed according to the STARD guidelines,[25] as documented in the relevant publications.[16–21,23]

2. Derivation cohort

The derivation cohort consisted of 1,572 consecutive symptomatic patients referred to colonoscopy in Ourense, Spain, between March 2012 and September 2013, who were included in the derivation cohort of the COLONPREDICT study.[23] The exclusion criteria have been detailed elsewhere.[23]

All patients collected one faecal sample from a single bowel movement during the week before the colonoscopy. They were specifically instructed to sample the passed faeces where no blood was visible. f-Hb was determined using the automated OC-SENSOR MICRO analyser (Eiken Chemical Co., Ltd, Tokyo, Japan). Colonoscopy was performed blind to the analytical results. The characteristics of this derivation cohort and the main endoscopic findings are shown in Table 1.

3. Validation cohort:

The validation cohort included 3,976 symptomatic patients recruited in five studies evaluating the diagnostic accuracy of different FIT analytical systems for CRC, advanced neoplasia (AN), and significant colonic lesion (SCL) detection or exclusion in symptomatic patients.[17,19–21,23] The characteristics of the different cohorts with respect to age, sex, colonoscopy findings, reasons for primary care referral and FIT system used are shown in Table 1. The particulars of these five studies are documented in detail in the relevant publications. [17-21]

All studies were approved by the local Clinical Research Ethics Committees and patients provided written informed consent in the Spanish studies. In the three Scottish studies, Research Ethics Committees stated that, if the patient returned a sample for f-Hb measurement, consent was implied. Estimates of f-Hb were quantitated as µg Hb/g faeces so that results could be compared across analytical systems.[26]

4. **Main outcome:**

The main outcome of this study was CRC detection. The secondary outcomes were advanced neoplasia (AN) and significant colonic lesions (SCL) diagnosis. AN was defined in the Spanish studies as advanced adenomas (≥10 mm, villous histology, high-grade dysplasia) or CRC and, in the studies done in Scotland, more simply as higher-risk adenoma (>10 mm, or more than three) or CRC. SCL was defined in Spain as any of the following: CRC, AN, polyposis (>10 polyps of any histology, including serrated lesions), colitis (any aetiology), polyps ≥10 mm (including serrated lesions), complicated diverticular disease (diverticulitis, bleeding), colonic ulcer or bleeding angiodysplasia, In Scotland,

again more simply, SCL was defined as any of CRC, AN or inflammatory bowel disease (IBD – Crohn's or ulcerative colitis); other lesions (non-advanced adenomas, non-complicated diverticular disease, polyps <10mm, non-bleeding angiodysplasia, haemorrhoids) were considered non-significant colonic lesions.

5. **Development of the prediction tool**

Three variables were included in the FAST Score prediction tool: the acronym is based on the **F**aecal haemoglobin concentration, **A**ge and **S**ex **T**est Score. Before logistic regression, we performed a univariate analysis using Generalised Additive Model models with smoothing splines for continuous variables. The objective of this analysis was to determine, in those nonlinear variables, the different strata or classes. Age was introduced as a continuous variable. In contrast, f-Hb had to be introduced as a categorical variable since it is not appropriate for introduction as a continuous variable in a logistic regression model. This is because f-Hb did not have a normal distribution, even after logarithmic transformation; in addition, the risk of CRC did not have a linear relationship to f-Hb in spite of f-Hb being related to colonic disease severity. On account of previous findings,[16–21] we decided to use three thresholds in f-Hb concentration: 0 µg Hb/g faeces (with the theoretical potential to rule out CRC and SCL),[19,21] 20 µg Hb/g faeces (the threshold used in many CRC screening programmes) and, finally, 200 µg Hb/g faeces, the upper limit of the analytical working range of the most commonly used FIT analytical system, the OC-Sensor. (Eiken) In consequence four categories were defined: (1) 0 µg Hb/g faeces, (2) between 0 and below 20 µg Hb/g faeces, (3) between 20 and

below 200 µg Hb/g faeces and (4) more than 200 µg Hb/g faeces. The regression coefficients were used to construct a CRC prediction Score, where the dependent variable was the presence or absence of CRC. We calculated the area under the curve (AUC) from the receiver operating characteristic (ROC) curve analyses. The calibration of the model in the derivation cohort was evaluated with the Hosmer-Lemeshow test

In order to evaluate the diagnostic yield of the final prediction tool, two threshold Scores were established with 90% and 99% sensitivity for CRC detection, and the diagnostic accuracy for CRC, AN and SCL were determined at each of these thresholds. Using these threshold Scores, the cohorts were dissected into three groups: high, intermediate and low risk for CRC detection. The number of patients, the positive predictive value (PPV) and the number needed to scope (NNS) to detect one CRC, AN, or SCL were calculated for each group. Finally, we expressed the differences in risk as Odds Ratios (OR) with 95% confidence intervals (CI).

6. **External validation:**

We used the coefficients to calculate the FAST score for each patient in the validation dataset. Those patients that met the criteria for 90% and 99% sensitivity were determined. The diagnostic accuracy of the model in the derivation and the validation cohorts were compared with, firstly, with the AUC derived from ROC curves and, secondly, with the Chi-square test to determine differences in sensitivity and specificity at the two threshold Scores between the cohorts for CRC, AN and SCL detection. The calibration of the model in the validation cohort was evaluated with the Hosmer-Lemeshow test.

7. **Diagnostic accuracy according to additional variables:**

Finally, a post-hoc analysis of the model was performed in the validation cohort to determine if its diagnostic accuracy for CRC detection was altered on the basis of the level of healthcare at which the patient was referred for colonoscopy (primary or secondary), the country in which the study was performed (Spain or Scotland), the CRC prevalence (<5% or >5%) in the study group, sex (male or female), age (<50 or ≥50 years), the individual study and, finally, the particular FIT analytical system used to estimate f-Hb. In order to perform this analysis, diagnostic accuracy was compared with AUC derived from ROC curves and sensitivity and specificity with the Chi-square test.

Data are reported with 95% CI. A p-value <0.05 was considered to be statistically significant. Analysis was carried out using SPSS statistical software, version 15.0 (SPSS Inc., Chicago, IL) and EPIDAT 3.1 (Dirección Xeral de Saúde Pública, Santiago de Compostela, Spain).

RESULTS

1. Development of the Score prediction model

The odds ratio (OR) of the variables included in the model were: age (years): 1.03, 95% (CI 1.02-1.05), male sex: 1.6 (95% CI: 1.1-2.3) and f-Hb (0-<20 µg Hb/g faeces): 2.0 (95% CI: 0.7-5.5), (20-<200 µg Hb/g): 16.8 (95% CI: 6.6-42.0), ≥200 µg Hb/g: 65.7 (95% CI: 26.3-164.1). The mathematical formula to calculate the FAST Score (b coefficient) is as follows: 0.684 x f-Hb (0, 20) µg Hb/g faeces + 2.824 x f-Hb [20, 200) µg Hb/g faeces + 4.184 x f-Hb ≥200 µg Hb/g faeces + 0.031 x age (years) + 0.479 x sex (male). The intercept term of the equation is -6.689. The FAST Score had an AUC for CRC diagnosis in the ROC analysis of 0.88 (95% CI: 0.85-0.90). The Hosmer Lemeshow test significance was p=0.4.

2. Diagnostic accuracy of the Score in the derivation cohort

The thresholds for the b-coefficient of the FAST Score with 90 and 99% sensitivity were 4.50 and 2.12, respectively. The b-coefficient was at least 4.50 in 37.1% and 2.12 in 88.0% of the patients included in the derivation cohort. Data on the diagnostic accuracy of the Score for CRC detection using these two threshold Scores are shown in Table 2. The overall AUC for AN detection was 0.82 (95% CI: 0.80-0.84) and for SCL was 0.82 (0.79-0.84). At the threshold Score of 4.50, the tool had a sensitivity for AN and SCL detection of 75.4% (95% CI: 70.9-79.4) and 72.7% (95% CI: 68.4-76.7), respectively. Moreover, at the threshold Score of 2.12, the sensitivity rose to 98.8% (95% CI: 97.1-99.6) and 97.8% (95% CI 95.9-98.9) for AN and SCL, respectively.

According to these two threshold Scores, the derivation cohort was dissected into three risk groups: high (Score ≥4.50), intermediate (Score <4.50 and ≥2.12) and low (Score <2.12). The diagnostic yields of this classification for CRC, AN and SCL detection are shown in Table 3. In summary, while the NNS to detect a CRC, or an AN, or a SCL were 35.7, 8.1, and 6.9 in the intermediate-risk group, the NNS in the high-risk group were 3.0, 1.8 and 1.7, respectively. No CRC was detected in the low-risk group and the NNS to detect an AN, or a SCL in this group rose to 37.0 and 18.9. The OR in the high-risk group for CRC detection was 17.5 (95% CI: 11.1-27.8) when compared with the intermediate-risk group.

3. Validation of the Score in the validation cohort

In the ROC analysis, the FAST Score was more accurate for CRC detection in the validation cohort (AUC: 0.91, 95% CI: 0.90-0.93; p = 0.005) as shown in Figure 1. This difference in the diagnostic accuracy was related to a statistically significant increase in specificity in the validation cohort as shown in Table 4. In contrast, there were no differences in sensitivity for CRC detection between the two cohorts. With respect to the calibration in the validation cohort, the Hosmer Lemeshow test significance was p=0.01. In addition, the diagnostic accuracy for AN (AUC: 0.79, 95% CI: 0.77-0.81; p = 0.05) and SCL (AUC: 0.78, 95% CI: 0.76-0.80; p = 0.01) was reduced in the validation cohort. This reduction in overall accuracy was especially relevant in the Scottish studies: AUC for AN: 0.75 (95% 0.70-0.79; p = 0.002) and for SCL: 0.75 (95%0.71-0.79; p < 0.003). In contrast, the differences in the overall accuracy in the studies performed in Spain were statistically non-significant: AUC for AN: 0.81 (95% CI

0.78-0.83; p = 0.2) and for SCL: 0.79 (95% 0.77-0.81; p = 0.08). Although the specificity for these two groups of colonic disease was also significantly increased in the validation cohort, the sensitivity was also reduced. These differences in sensitivity and specificity also occurred when comparing the data from the Scottish and Spanish studies (Supplementary Table 1).

The FAST Score also discriminates the risk of CRC, AN and SCL detection in the validation cohort as shown in Table 3. In this regard, 851 patients (21.4%) met high-risk group criteria with a 21.7% PPV for CRC detection; 2,378 (59.8%) met intermediate-risk group criteria with a 0.9% PPV and, finally, 747 (18.8%) patients met low-risk group criteria with a 0.0% PPV for CRC detection in the validation cohort.

4. Diagnostic accuracy in the validation cohort according to the characteristics of the groups.

In the post-hoc analysis, possible differences in the diagnostic accuracy of the FAST Score for CRC detection were assessed in the validation cohort according to several variables, as shown in Table 5. No significant differences in sensitivity for CRC detection were detected in any of the groups assessed. In addition, no differences were found in the diagnostic accuracy when comparing the patients referred for colonoscopy from primary care with those referred from secondary care. In contrast, the diagnostic accuracy, either overall or as specificity, was increased in the groups with a lower CRC prevalence. In this respect, differences were found in the CRC prevalence according to sex (male: 7.1%, female: 3.6%, p < 0.001), age (<50 years: 1.3%, ≥50 years: 6.1%; p < 0.001) and country (Spain: 6.5%, Scotland: 3.0%, p<0.001). In contrast, the

CRC prevalence was similar in both healthcare levels referring patients (primary: 5.6%, secondary: 4.9%, p = 0.4). Finally, the diagnostic accuracy was compared between the most commonly used analytical system (OC-Sensor) and the other analytical methods used, including the widely used HM-JACKarc, and between the validation cohort of the COLONPREDICT study and the rest of the study groups. As shown in Table 5, diagnostic accuracy was equal or higher with respect to the validation cohort, both for the analytical system and the study groups.

DISCUSSION

1. Statement of principal findings

Because f-Hb is dependent on age and sex,[27] we developed and assessed the diagnostic accuracy of an easy to calculate prediction tool based not only on f-Hb, as originally suggested by Chen et al,[28] but also age and sex.[20] This prediction tool, termed the FAST Score, is highly accurate for CRC detection in symptomatic patients and allows establishment of three risk groups: CRC can be ruled out in those in the lowest risk group. In addition, we have confirmed that the threshold Scores we have examined are equally sensitive for CRC irrespective of, country, CRC prevalence, age, sex, healthcare level and analytical system used to estimate f-Hb. In contrast, this prediction tool was less accurate for AN and SCL detection in the validation cohort.

2. Strengths and weaknesses of the study

We had the opportunity to access the individual data of the patients included in five of the six studies that have evaluated different analytical systems for the measurement of f-Hb in symptomatic patients referred for colonoscopy.[16–21] These studies have evaluated the application of FIT in a wide variety of clinical settings with differences in the strategy for invitation to the study, the analytical systems, the referral criteria to colonoscopy and the epidemiological characteristics of the cohorts included (age, sex and CRC, AN and SCL prevalence).. On the basis of these data, we have developed and externally validated a multivariable prediction tool according to the TRIPOD guidelines.[24] A CRC prediction model, COLONPREDICT, had been previously developed and

validated.[23] This prediction model was compared with the NICE referral guidelines used at that time[3] demonstrating that a prediction model using f-Hb and other variables was more accurate than symptom-based referral criteria.[23] In the case of the FAST Score, its greatest advantage is simplicity and it could be calculated automatically by laboratory information management systems (LIMS) widely used to report results in laboratory medicine, with or without appropriate interpretative comments added by the LIMS, or through simple calculators developed for the Internet or as easy to download apps, since age and sex are generally available in health care systems and, if not, are easily collected from patients with symptoms prior to referral for colonoscopy along with the f-Hb. We think the simplicity of this prediction model may encourage the implementation of the FAST Score in comparison with previous developed prediction models for CRC detection that are based on several clinical and analytical variables which must be collected from disparate sources before the model is applied,[12-15]

The FAST Score allows the rapid allocation of a patient to one of three risk groups: a high risk group where 90% of CRC are detected and these might be referred on an urgent basis, an intermediate risk group where 9% additional CRC are detected and such patients might be assessed further before referral, and a low risk group where, hypothetically, the remaining 1% would be found. This low risk group, with a negligible PPV accounts for nearly 19% of the validation cohort and these patients would probably not benefit from colonoscopy or possibly any further examination. In contrast, the NICE referral criteria for suspected CRC have a 68.2% sensitivity and a 50.2% specificity for CRC detection, and, in reality, these criteria cannot rule out CRC.[23]

The main weakness of our study is that the patients included in these studies had been selected a priori by health care professionals for further endoscopic evaluation. In consequence, we believe that the diagnostic accuracy of the FAST Score that we have developed, should be externally evaluated in a population with gastrointestinal symptoms attending primary care before its use is widely adopted. In fact, our findings strongly suggest that the FAST Score diagnostic accuracy will increase in populations in which CRC is at a low prevalence since this will result in an increase in specificity. We also propose that our f-Hb, age and sex based prediction tool should be compared within a randomised controlled trial with the currently available referral guidelines, such as the recent recommendations from NICE[11] in terms of sensitivity, efficiency (endoscopy referrals to detect a CRC, delays) and, finally, effect on CRC survival.[29] Additionally, we have found significant differences in the calibration of the model in the validation cohort, probably related to the differences in the CRC prevalence between the derivation (13.7%) and the validation cohort (5.2%). These differences do not reduce the value of our findings. The prediction model is not aimed at estimating the individual risk but, rather, to determine different diagnostic strategies in symptomatic patients. However, the FAST Score will require not only validation in primary healthcare but also recalibration in this setting before using it to provide risk-adjusted outcomes.

Another limitation is the reduced diagnostic accuracy for detection of additional SCL. In part, this might be related to the different definitions used in the studies for AN and SCL in Spain and in Scotland. On the other hand, although this reduced accuracy could be considered as a limitation, there are

several arguments to support use of the FAST Score in this broader clinical perspective. Firstly, there are very few data published on AN detection models and tools,[12,20] but there is no model or tool published to date for SCL detection, apart from the more complex COLONPREDICT score. As is widely recognised, there are criteria for the appropriateness of colonoscopy referral. These criteria have been evaluated in a prospective fashion showing that their diagnostic accuracy for both CRC and SCL detection is limited. As one example, a meta-analysis has shown that the diagnostic OR of the American Society for Gastrointestinal Endoscopy and European Panel on the Appropriateness of Gastrointestinal Endoscopy (EPAGE) criteria for relevant findings was 2.5 (95% CI: 1.2-5.6),[30] inferior to the OR of the FAST Score in both derivation and the validation cohorts. Further, although we cannot exclude relevant findings in the low risk group (PPV: 5.2-5.3%), the PPV in the inappropriate colonoscopy group according to the EPAGE II criteria ranges between 9.5 and 24.5%.[31,32] Finally, it is quite clear that the main endpoint of endoscopic evaluation of symptomatic patients must be the detection, or perhaps more importantly, the exclusion of those clinical conditions that threaten or significantly affect the quality of life, especially on account of the limited endoscopic resources available in many countries.

3. **Strengths and weaknesses in relation to other studies, discussing important differences in results**

The FAST Score was derived from the data from the same cohort as the COLONPREDICT score but, as commented previously, we had the opportunity to validate it in a wider geographical cohort. Both scores are tools based mainly

on f-Hb. In the case of the FAST score, we applied f-Hb in four categories, one of them 0 µg Hb/g faeces. Previous studies have shown the probability of ruling out CRC in patients with low f-Hb (<10 µg Hb/g faeces)[20,25] or f-Hb of 0 µg Hb/g faeces.[17] However, although this was not the case in our study, patients with this f-Hb had a 0.5, 0.06 and 0.015-fold risk when compared with those patients with 0-20, 20-200 and ≥200 µg Hb/ g faeces, respectively.

In comparison with the COLONPREDICT score, the FAST Score has lower diagnostic accuracy. The COLONPREDICT score detects 90 and 99% CRC in 30.9% and 60.5% of the derivation cohort, in contrast with the FAST Score, that requires 37.1% and 88.0% of the cohort to achieve these levels. However, these data should be further evaluated in real routine practice. Unlike the FAST score, COLONPREDICT requires an anorectal examination, tests done on venous blood and a detailed history to be performed before its calculation. So, the increased accuracy of the COLONPREDICT score in a controlled setting may be limited in daily practice because of its complexity compared to the simplicity of the FAST score that only requires f-Hb, age and sex. That is why we think both approaches should be compared in a primary care setting in a future study.

4. **Meaning of the study: possible explanations and implications for clinicians and policymakers**

Even though we now need to evaluate the applicability of the FAST Score in a routine primary care setting, the implications for clinicians and policymakers are quite clear. Prediction tools based on f-Hb measured by high quality automated quantitative FIT should be the cornerstone of the evaluation of the

symptomatic patient and used as the criterion to determine the appropriateness of the referrals for colonoscopy. We consider that the FAST Score does provide an objective tool to guide who requires further investigation in secondary care. In fact, recently published (2015) NICE referral guidelines for suspected cancer have included the use of tests for the presence of occult blood in faeces in symptomatic patients with a PPV for CRC detection below 3%.[11] However, we think that f-Hb-based prediction tools could be a better basis for the evaluation of all symptomatic patients.

Furthermore, we believe strategies for prompt diagnosis of patients with suspected CRC should be designed based on sensitivity rather than on the PPV. As an example, if we use the 3% PPV threshold for CRC for further evaluation determined in the 2015 NICE guidelines,[11] we would miss one out of ten CRC, since the PPV in the FAST intermediate risk group in the derivation and validation cohorts is 2.8% and 0.9%, respectively. Thus, we consider that prompt diagnosis strategies should determine a 90% sensitivity threshold for urgent referral and a 99% sensitivity referral threshold for colonoscopy evaluation. The secondary endpoint should be specificity as long as low specificity is associated with unnecessary referrals to colonoscopy, with its corresponding risks. In fact, the specificity of the FAST Score at the 90% sensitivity threshold seems acceptable when compared with other available clinical criteria, e.g., NICE referral criteria.[18] In contrast, the specificity at the 99% sensitivity threshold is extremely low. However, it must be emphasised that, at this threshold, to the Score rules out CRC.

5. **Unanswered questions and future research**

Two main issues need to be answered in the future. As stated previously, the diagnostic accuracy and applicability of the FAST Score tool in a primary care setting must be addressed in a prospective study and, ideally, compared with the COLONPREDICT score and current age and symptom-based referral guidelines. Secondly, further prediction tools based on laboratory findings other than f-Hb should be designed and evaluated in a primary care setting. In this respect, the evaluation of newer CRC biomarkers in risk scoring models would be a necessary prerequisite to their introduction as investigations to assist the CRC diagnosis process in symptomatic patients.

References

1. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, Alvarado M, Anderson HR, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012;380:2095–128.

2. Atkin WS, Valori R, Kuipers EJ, Hoff G, Senore C, Segnan N, Jover R, Schmiegel W, Lambert R, Pox C, International Agency for Research on Cancer. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Colonoscopic surveillance following adenoma removal. *Endoscopy* 2012;44 Suppl 3:SE151-63.

3. NICE Clinical Guideline 27- Referral guidelines for suspected cancer. April 2011. www.nice.org.uk/CG027. 2011.

4. Hamilton W. Five misconceptions in cancer diagnosis. *Br J Gen Pract* 2009;59:441–5, 447; discussion 446.

5. Scottish Intercollegiate Guidelines Network (SIGN). Diagnosis and management of colorectal cancer. Edinburgh: SIGN; 2011. (SIGN publication no. 126). [December 2011]. Available from URL: http://www.sign.ac.uk.

6. Vega-Villaamil P, Salve-Bouzo M, Cubiella J, Valentín-Gómez F, Sánchez-Hernández E, Gómez-Fernández I, Fernández-Seara J. Evaluation of the implementation of Galician Health Service indications and priority levels for colonoscopy in symptomatic patients: prospective, cross-sectional study. *Rev Esp Enferm Dig* 105:600–8.

7. Mansouri D, McMillan DC, Crearie C, Morrison DS, Crighton EM, Horgan PG. Temporal trends in mode, site and stage of presentation with the introduction of colorectal cancer screening: a decade of experience from the West of Scotland. *Br J Cancer* 2015;113:556–61.

8. Ford AC, Veldhuyzen van Zanten SJO, Rodgers CC, Talley NJ, Vakil NB, Moayyedi P. Diagnostic utility of alarm features for colorectal cancer: systematic review and meta-analysis. *Gut* 2008;57:1545–53.

9. Jellema P, van der Windt D a. WM, Bruinvels DJ, Mallen CD, van Weyenberg SJB, Mulder CJ, de Vet HCW. Value of symptoms and additional diagnostic tests for colorectal cancer in primary care: systematic review and meta-analysis. *BMJ* 2010;340:c1269.

10. Astin M, Griffin T, Neal RD, Rose P, Hamilton W. The diagnostic value of symptoms for colorectal cancer in primary care: a systematic review. *Br J Gen Pract* 2011;61:e231-43.

11. Cancer NCC for. Suspected cancer : recognition and management of suspected cancer in children, young people and aduls [Internet]. 2015 [cited 2015 Aug 20];Appendices A-J4. Available from:

http://www.nice.org.uk/guidance/ng12

12. Williams TGS, Cubiella J, Griffin SJ, Walter FM, Usher-Smith JA. Risk prediction models for colorectal cancer in people with symptoms: a systematic review. *BMC Gastroenterol* 2016;16:63.

13. Hodder RJ, Ballal M, Selvachandran SN, Cade D. Pitfalls in the construction of cancer guidelines demonstrated by the analyses of colorectal referrals. *Ann R Coll Surg Engl* 2005;87:419–26.

14. Rai S, Ballal M, Thomas WM, Miller a S, Jameson JS, Steward WP. Assessment of a patient consultation questionnaire-based scoring system for stratification of outpatient risk of colorectal cancer. *Br J Surg* 2008;95:369–74.

15. Ballal MS, Selvachandran SN, Maw A. Use of a patient consultation questionnaire and weighted numerical scoring system for the prediction of colorectal cancer and other colorectal pathology in symptomatic patients: a prospective cohort validation study of a Welsh population. *Colorectal Dis* 2010;12:407–14.

16. Oono Y, Iriguchi Y, Doi Y, Tomino Y, Kishi D, Oda J, Takayanagi S, Mizutani M, Fujisaki T, Yamamura A, Hosoi T, Taguchi H, et al. A retrospective study of immunochemical fecal occult blood testing for colorectal cancer detection. *Clin Chim Acta* 2010;411:802–5.

17. McDonald PJ, Digby J, Innes C, Strachan J a, Carey F a, Steele RJC, Fraser CG. Low faecal haemoglobin concentration potentially rules out significant colorectal disease. *Colorectal Dis* 2013;15:e151-9.

18. Cubiella J, Salve M, Díaz-Ondina M, Vega P, Alves MT, Iglesias F, Sánchez E, Macía P, Blanco I, Bujanda L, Fernández-Seara J. Diagnostic accuracy of the faecal immunochemical test for colorectal cancer in symptomatic patients: comparison with NICE and SIGN referral criteria. *Colorectal Dis* 2014;16:O273-82.

19. Mowat C, Digby J, Strachan JA, Wilson R, Carey FA, Fraser CG, Steele RJC. Faecal haemoglobin and faecal calprotectin as indicators of bowel disease in patients presenting to primary care with bowel symptoms. *Gut* 2015;1–7.

20. Rodríguez-Alonso L, Rodríguez-Moranta F, Ruiz-Cerulla A, Lobatón T, Arajol C, Binefa G, Moreno V, Guardiola J. An urgent referral strategy for symptomatic patients with suspected colorectal cancer based on a quantitative immunochemical faecal occult blood test. *Dig Liver Dis* 2015;47:797–804.

21. Godber IM, Todd LM, Fraser CG, MacDonald LR, Younes H Ben. Use of a faecal immunochemical test for haemoglobin can aid in the investigation of patients with lower abdominal symptoms. *Clin Chem Lab Med* 2015;

22. Usher-Smith JA, Walter FM, Emery J, Win AK, Griffin SJ. Risk prediction models for colorectal cancer: a systematic review. *Cancer Prev Res (Phila)* 2015;

23. Cubiella J, Vega P, Salve M, Díaz-Ondina M, Alves MT, Quintero E,

Álvarez-Sánchez V, Fernández-Bañares F, Boadas J, Campo R, Bujanda L, Clofent J, et al. Development and external validation of a faecal immunochemical test-based prediction model for colorectal cancer detection in symptomatic patients. *BMC Med* 2016;14:128.

24. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med* 2015;162:55–63.

25. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis C a, Glasziou PP, Irwig LM, Lijmer JG, Moher D, Rennie D, de Vet HCW. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Fam Pract* 2004;21:4–10.

26. Fraser CG, Allison JE, Halloran SP, Young GP. A proposal to standardize reporting units for fecal immunochemical tests for hemoglobin. *J Natl Cancer Inst* 2012;104:810–4.

27. Fraser CG, Rubeca T, Rapi S, Chen LS, Chen HH. Faecal haemoglobin concentrations vary with sex and age, but data are not transferable across geography for colorectal cancer screening. *Clin Chem Lab Med* 2014;52:1211–6.

28. Chen L-S, Yen AM-F, Chiu SY-H, Liao C-S, Chen H-H. Baseline faecal occult blood concentration as a predictor of incident colorectal neoplasia: longitudinal follow-up of a Taiwanese population-based colorectal cancer screening cohort. *Lancet Oncol* 2011;12:551–8.

29. Sackett DL, Haynes RB. The architecture of diagnostic research. *BMJ* 2002;324:539–41.

30. Hassan C, Di Giulio E, Marmo R, Zullo A, Annibale B. Appropriateness of the indication for colonoscopy: systematic review and meta-analysis. *J Gastrointestin Liver Dis* 2011;20:279–86.

31. Gimeno García AZ, González Y, Quintero E, Nicolás-Pérez D, Adrián Z, Romero R, Alarcón Fernández O, Hernández M, Carrillo M, Felipe V, Díaz J, Ramos L, et al. Clinical validation of the European Panel on the Appropriateness of Gastrointestinal Endoscopy (EPAGE) II criteria in an open-access unit: a prospective study. *Endoscopy* 2012;44:32–7.

32. Eskeland SL, Dalén E, Sponheim J, Lind E, Brunborg C, de Lange T. European panel on the appropriateness of gastrointestinal endoscopy II guidelines help in selecting and prioritizing patients referred to colonoscopy--a quality control study. *Scand J Gastroenterol* 2014;49:492–500.

Figure legends

Figure 1:

ROC curves for the FAST Score for colorectal cancer detection in derivation and validation cohorts. The AUC of the ROC curves are compared with the Chi-square homogeneity area test.

ROC: receiver operating characteristic, AUC: area under the curve, 95% CI: 95% confidence interval.

Table 1: Characteristics of the different cohorts included in the study.

Authors, date, country, reference	Number	Male sex	Age in years (median and range)	Colorectal cancer	Advanced neoplasia	Significant colorectal lesions	Rectal bleeding	Change in bowel habit	Primary health care referral	Analytical system for estimation of faecal haemoglobin concentration
Derivation cohort										
Cubiella et al, 2016, Spain.[23]	1,572	51.5%	68 (20-96)	13.7%	26.7%	29.5%	59.5%	57.2%	22.9%	OC-Sensor: 100%
Validation cohorts										
McDonald et al, 2013, Scotland[,17]	280	40.4%	60 (15-89)	2.1%	12.1%	21.6%	25.7%	6.1%	100.0%	OC-Sensor: 100%
Mowat et al, 2015, Scotland[,19]	750	45.5%	64 (16-90)	3.7%	10.1%	14.7%	33.9%	42.7%	100.0%	OC-Sensor: 100%
Godber et al, 2016, Scotland[,21]	484	42.1%	63 (18-84)	2.3%	6.2%	9.3%	24.0%	27.5%	0.0%	HM-JACKarc:100%
Rodriguez-Alonso et al, 2015, Spain, [20]	1,003	46.9%	63 (18-90)	3.0%	13.3%	15.5%	34.2%	35.6%	17.2%	OC-Sensor: 100%
Cubiella et al, 2016, Spain,[23]	1,459	48.7%	64 (19-100)	9.0%	21.5%	29.2%	51.5%	52.7%	39.2%	OC-Sensor: 49.7% OC-Auto 3 Latex: 13.8% FOB Gold: 2.4% Linear i-FOB: 34.1%
Overall	3,976	46.2%	65 (15-100)	5.2%	14.8%	19.5%	38.6%	40.1%	37.6%	OC-Sensor: 69.3% OC-Auto 3 Latex: 5.1% FOB Gold: 0.9% Linear i-FOB: 12.5% HM-JACKarc:12.2%

Table 2: Diagnostic accuracy of the FAST Score at two threshold Scores with 90% and 99% sensitivity for colorectal cancer (CRC), advanced neoplasia (AN) and significant colonic lesion (SCL) detection in the derivation cohort

	FAST Score ≥4.50			**FAST Score ≥2.12**		
	CRC	**AN[1]**	**SCL[2]**	**CRC**	**AN[1]**	**SCL[2]**
Number positive	37.1%			88.0%		
Sensitivity[3]	89.8% (84.7-93.3)	75.4% (70.9-79.4)	72.7% (68.4-76.7)	100.0% (97.8-100.0)	98.8% (97.1-99.6)	97.8% (95.9-98.9)
Specificity[3]	71.3% (68.8-73.7)	76.9% (74.3-79.3)	77.8% (75.2-80.2)	13.9% (12.1-15.9)	15.9% (13.9-18.2)	16.1% (14.0-18.4)
Positive predictive value[3]	33.2% (29.4-37.2)	54.4% (50.2-58.5)	57.8% (53.7-61.9)	15.6% (13.7-17.6)	30.0% (27.6-32.5)	32.8% (30.3-35.3)
Negative predictive value[3]	97.8% (96.6-98.6)	89.6% (87.4-91.4)	87.2% (84.9-89.2)	100% (97.5-100.0)	97.3% (93.5-99)	94.7% (90.2-97.3)
Positive likelihood ratio[4]	3.13 (2.84-3.44)	3.27 (2.90-3.68)	3.28 (2.90-3.71)	1.16 (1.14-1.19)	1.18 (1.14-1.21)	1.17 (1.13-1.2)
Negative likelihood ratio[4]	0.14 (0.10-0.21)	0.32 (0.27-0.38)	0.35 (0.30-0.41)	NE	0.07 (0.03-0.18)	0.13 (0.07-0.25)
Diagnostic odds ratio[4]	21.8 (13.8-34.4)	10.2 (7.8-13.2)	9.4 (7.3-12.0)	NE	15.7 (6.4-38.4)	8.7 (4.5-16.5)

[1]Advanced neoplasia = colorectal cancer plus advanced adenoma (see text)

[2]Significant colonic lesion = colorectal cancer plus advanced adenoma (see text) plus other significant pathology (see text)

[3]Values are expressed as percentages with 95% confidence intervals in parentheses.

[4]Values are expressed as absolute numbers and 95% confidence intervals in parentheses.

NE: not evaluable.

Table 3: Diagnostic yield for colorectal cancer, advanced neoplasia and significant colonic lesion detection according to the FAST Score risk classification in the derivation cohort and validation cohort

		Derivation cohort (1,572)			Validation cohort (3,976)		
		Low[1]	**Intermediate[2]**	**High[3]**	**Low[1]**	**Intermediate[2]**	**High[3]**
Number patients (%)		12	50.9	37.1	18.8	59.8	21.4
Colorectal cancer	**PPV (%)**	0	2.8	33.2	0	0.9	21.7
	NNS	NE	35.7	3.0	NE	111.1	4.6
	OR (95% CI)	NE[6]	17.5 (11.1-27.8) [7]		NE[6]	29.4 (18.9-45.4) [7]	
		NE[8]			NE[8]		
Advanced neoplasia[4]	**PPV (%)**	2.7	12.3	54.4	2.6	8.9	41.7
	NNS	37	8.1	1.8	38.5	11.2	2.4
	OR (95% CI)	5.1 (2.1-12.7) [6]	8.5 (6.5-11.1) [7]		3.7 (2.3-6) [6]	7.3 (6.1-9) [7]	
		43.6 (17.7-107) [8]			27.4 (17-44) [8]		
Significant colonic lesion[5]	**PPV (%)**	5.3	14.5	57.8	5.6	11.9	52.6
	NNS	18.9	6.9	1.7	17.8	8.4	1.9
	OR (95% CI)	3.0 (1.5-6.0) [6]	8.1 (6.2-10.4) [7]		2.2 (1.6-3.2) [6]	8.2 (6.8-9.8) [7]	
		24.4 (12.6-47.1) [8]			18.5 (13.2-26) [8]		

[1]Low-risk cohort: FAST Score <2.12.

[2]Intermediate-risk cohort: FAST Score ≥2.12 and <4.50.

[3]High-risk cohort: FAST Score ≥4.50.

[4]Advanced neoplasia = colorectal cancer plus advanced adenoma (see text)

[5]Significant colonic lesion = colorectal cancer plus advanced adenoma plus other significant pathology (see text)

[6]OR and 95% CI when the intermediate risk compared with the low-risk group

[7]OR and 95% CI when the high risk compared with the intermediate-risk group.

[8]OR and 95% CI when the high risk compared with the low-risk group.

PPV: positive predictive value, NE: non-evaluable, NNS: number needed to scope; OR, odds ratio; 95% CI, 95% confidence interval.

Table 4: Sensitivity and specificity for colorectal cancer (CRC), advanced neoplasia (AN) and significant colonic lesion (SCL) of the FAST Score at the thresholds (4.50 and 2.12) with 90% and 99% sensitivity for CRC detection in the derivation and validation cohorts

FAST Score		FAST Score ≥4.50		FAST Score ≥2.12	
		Sensitivity[1]	Specificity[1]	Sensitivity[1]	Specificity[1]
CRC	Derivation	89.8% (84.7-93.3)	71.3% (68.8-73.7)	100.0% (97.8-100)	13.9% (12.1-15.9)
	Validation	89.3% (84.1-93.0)	82.3% (81.1-83.5)	100.0% (97.7-100)	19.8% (18.6-21.1)
	Significance[4]	1	<0.001	NE	<0.001
AN[2]	Derivation	75.4% (70.9-79.4)	76.9% (74.3-79.3)	98.8% (97.1-99.6)	15.9% (13.9-18.2)
	Validation	60.7% (56.6-64.7)	85.4% (84.1-86.5)	96.7% (94.9-98)	21.5% (20.1-22.9)
	Significance[4]	<0.001	<0.001	0.03	<0.001
SCL[3]	Derivation	72.7% (68.4-76.7)	77.8% (75.2-80.2)	97.8% (95.9-98.9)	16.1% (14.0-18.4)
	Validation	57.8% (54.3-61.3)	87.4% (86.2-88.5)	94.5% (92.6-96)	22.0% (20.6-23.5)
	Significance[4]	<0.001	<0.001	0.005	<0.001

[1]Values are expressed as percentages with 95% confidence intervals in parentheses.

[2]Advanced neoplasia = colorectal cancer plus advanced adenoma (see text)

[3]Significant colonic lesion = colorectal cancer plus advanced adenoma (see text) plus other significant pathology (see text)

[4]Significance of the sensitivity and specificity differences between both cohorts in the Chi-square test. Differences with $p<0.05$ are considered statistically significant.

NE: non-evaluable.

Table 5: Diagnostic accuracy for colorectal cancer detection of the FAST Score in the validation cohort according to the characteristics of the cohorts and the patients included: AUC and sensitivity and specificity at the Score thresholds (4.50 and 2.12) with 90% and 99% sensitivity for CRC detection.

		FAST score		FAST Score ≥4.50				FAST Score ≥2.12	
		AUC (95% CI)	p[1]	Sensitivity	p[2]	Specificity	p[2]	Specificity	p[2]
Healthcare setting	**Primary (1,496)**	0.90 (0.87-0.93)	0.2	89.2%	1	81.8%	0.5	21.2%	0.1
	Secondary (2,480)	0.92 (0.91-0.94)		89.3%		82.6%		19.0%	
Region	**Spain (2,462)**	0.90 (0.89-0.92)	0.6	86.7%	0.5	85.4%	<0.001	18.3%	0.004
	Scotland (1,514)	0.92 (0.88-0.95)		89.3%		82.3%		22.2%	
Age	**<50 years (761)**	0.93 (0.90-0.96)	0.07	80.0%	0.3	90.6%	<0.001	56.0%	<0.001
	≥50 years (3,215)	0.91 (0.88-0.92)		89.8%		80.3%		10.8%	
Sex	**Male (1,838)**	0.89 (0.86-0.91)	0.01	90.0%	0.8	77.6%	<0.001	8.0%	<0.001
	Female (2,138)	0.93 (0.91-0.95)		88.2%		86.3%		29.6%	
Colorectal cancer prevalence	**<5% (2,517)**	0.92 (0.90-0.95)	0.03	88.0%	0.6	86.3%	<0.001	23.4%	<0.001
	≥5% (1,459)	0.88 (0.86-0.91)		90.1%		75.0%		13.2%	

FIT[3]	**OC-Sensor (2,758)**	0.92 (0.91-0.94)		90.1%		83.1%		23.7%	
	OC-Auto 3 Latex (202)	0.88 (0.81-0.94)	0.1	95.2%	0.7	56.4%	<0.001	7.2%	<0.001
	FOB Gold reagents (35)	0.88 (0.63-1.00)	0.7	100%	1	72.7%	0.1	24.2%	1
	Linear i-FOB (497)	0.89 (0.84-0.93)	0.1	82.9%	0.2	86.4%	0.08	8.4%	<0.001
	HM-JACKarc (484)	0.97 (0.94-1.00)	0.02	90.9%	1	84.6%	0.4	14.0%	<0.001
Study[4]	**Cubiella et al 2016 (1,459)**	0.88 (0.86-0.91)		90.1%		75%		13.2%	
	Godber et al 2015 (484)	0.97 (0.94-1.00)	<0.001	90.9%	1	84.6%	<0.001	14.0%	0.7
	Mowat et al 2015 (750)	0.89 (0.83-0.94)	0.9	82.1%	0.3	85.5%	<0.001	26.3%	<0.001
	McDonald et al 2013 (280)	0.95 (0.92-0.99)	<0.001	100%	1	86.9%	<0.001	25.5%	<0.001
	Rodriguez-Alonso et al 2015 (1,003)	0.93 (0.91-0.95)	0.003	90.0%	1	87.6%	<0.001	25.2%	<0.001

[1]Significance of the area under the curve differences between cohorts with the Chi-square homogeneity test. Differences with $p<0.05$ are considered statistically significant.

[2]Significance of the sensitivity and specificity differences between cohorts with the Chi-square test. Differences with $p<0.05$ are considered statistically significant.

[3]The cohorts are compared with the OC-Sensor cohort.

[4]The cohorts are compared with the Cubiella et al study cohort.

AUC, area under the curve; CI, confidence interval; FIT, faecal immunochemical test for haemoglobin

Supplementary table 1: Comparison of the diagnostic accuracy of the FAST Score for advanced neoplasia and significant colonic lesions in the Spanish and Scottish validation studies. Comparison with the derivation cohort.

Lesion	Thresholds of the FAST Score		Derivation	Scotland		Spain	
				Value	p[4]	Value	p[4]
Advanced neoplasia[1]	Sensitivity[3]	**4.50**	75.4% (70.9-79.4)	50.7% (42.2-59.2)	<0.001	63.9% (59.2-68.3)	<0.001
		2.12	98.8% (97.1-99.6)	95.0% (89.6-97.8)	0.01	97.3% (95.2-98.5)	0.1
	Specificity[3]	**4.50**	76.9% (74.3-79.3)	86.8% (84.8-88.5)	<0.001	84.4% (82.8-86.0)	<0.001
		2.12	15.9% (13.9-18.2)	23.2% (21.0-25.6)	<0.001	20.3% (18.6-22.1)	0.03
Significant colonic lesion[2]	Sensitivity[3]	**4.50**	72.7% (68.4-76.7)	50.4% (43.7-57.1)	<0.001	60.9% (56.6-65.0)	<0.001
		2.12	97.8% (95.9-98.9)	91.5% (86.9-94.7)	<0.001	95.8% (93.7-97.3)	0.07
	Specificity[3]	**4.50**	77.8% (75.2-80.2)	89.1% (87.3-90.8)	<0.001	86.2% (84.6-87.7)	<0.001
		2.12	16.1% (14-18.4)	23.8% (21.5-26.2)	<0.001	20.8% (19.0-22.7)	0.001

[1]Advanced neoplasia = colorectal cancer plus advanced adenoma (see text)

[2]Significant colonic lesion = colorectal cancer plus advanced adenoma (see text) plus other significant pathology (see text)

[3]Values are expressed as percentages with 95% confidence intervals in parentheses.

[4]Significance of the sensitivity and specificity differences between derivation cohort and the Scottish and Spanish studies included in the validation cohort with the Chi-square test. Differences with $p<0.05$ are considered statistically significant.

Supplementary note: Investigators of the COLONPREDICT study

Complexo Hospitalario Universitario de Ourense: Joaquín Cubiella, Pablo Vega, María Salve, Marta Díaz-Ondina, Irene Blanco, Pedro Macía, Eloy Sánchez, Javier Fernández-Seara.

Hospital Universitario de Canarias: Enrique Quintero, Natalia González-López

Complejo Hospitalario de Pontevedra: Victoria Álvarez Sánchez, José Mera, Juan Turnes.

Hospital Universitari Mútua de Terrassa: Fernando Fernández-Bañares, Victoria Gonzalo, Mar Pujals.

Registre del Càncer de Catalunya Pla Director d'Oncologia de Catalunya, Hospital Duran i Reynals, L'Hospitalet de Llobregat: Josepa Ribes, Ramón Cleries, Xavier Sanz.

Consorci Sanitari de Terrassa: Jaume Boadas, Sara Galter, Rebeca Gimeno, Eva Garcia-Lanuza

Corporació Sanitària i Universitària Parc Taulí: Rafel Campo, Marta Pujol, Eva Martínez-Bauer.

Departamento de Bioquímica, CATLAB, Viladecavalls, Barcelona: Antonio Alsius.

Donostia Hospital: Luis Bujanda, Jesús Bañales, María J Perugorria.

Hospital de Sagunto: Joan Clofent, Ana Garayoa.

Hospital Clínico Universitario de Zaragoza: Ángel Ferrández, Marina Solano Sánchez.

Hospital Dr. Josep Trueta: Leyanira Torrealba, Virginia Piñol.

Hospital Universitario de Móstoles: Daniel Rodriguez-Alcalde, Jorge López-Vicente.

Acknowledgements, competing interests, funding:

Sources of funding:

This study was funded by a grant from Instituto de Salud Carlos III (PI11/00094). JC has received an intensification grant through the European Commission funded "BIOCAPS" project (FP-7-REGPOT 2012-2013-1, Grant agreement no. FP7- 316265). The validation cohort recruitment was funded by a grant from Fundació de la Marató TV3 2012 (785/U/2013). JD was supported by a grant from Tenovus Tayside, Scotland. The funding institutions had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Competing interest declaration:

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: JC and MS had financial support from Instituto de Salud Carlos III for the submitted work but they had no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, and no other relationships or activities that could appear to have influenced the submitted work. CGF has undertaken consultancy with Immunostics, Ocean, New Jersey, USA; and Kyowa-Medex Co., Ltd, Tokyo, Japan: and has received travel support from Alpha Labs, Eastleigh, UK. The remaining authors had no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years and no other relationships or activities that could appear to have influenced the submitted work.

Copyright for authors

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) exploit all subsidiary rights of the Contribution, v) include electronic links from the Contribution to third-party material wherever it may be located and, vi) licence any third party to do any or all of the above.

Transparency: The lead author affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Data sharing: Input data for the model are available from the corresponding author on request.

Ethical approval: The ethical approval for each of the studies included are documented in detail in the relevant publications cited in the text.

Meeting presentations: This research was presented in the World Endoscopy Organization Colorectal Cancer Screening Committee Meeting held in Barcelona, 23rd October 2015; the 23rd United European Gastroenterology Week, held in Barcelona, 24-28th October 2015 and in the XIX Reunión Nacional de la Asociación Española de Gastroenterologia held in Madrid 1-4th March, 2016.

Details of contributors:

The authors' contributions were as follows: JC, JD, FRM, RJCS and CGF participated in the model development design; JD, LRA, PV, MS, MDO, JAS, CM, PJMcD, FAC, IMG, HBY, FRM, EQ, VAS, FFB, JB, RC, LB, AG, AF, VP, DRA and JG in the recruitment of the derivation and validation cohorts; JC, JD, FRM, RJCS and CGF in the derivation and validation of the prediction model; JC, JD, RJCS and CGF in the preparation of drafts of the manuscript. All authors contributed to the writing of the paper. All authors had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis. JC acted as full guarantor of the research.

Acknowledgements:

Ciberehd is funded by Instituto de Salud Carlos III.

We acknowledge María Ángeles López-Martínez and Manuel Rubio for their technical support in the inclusion of the validation cohort.

Provenance and peer review Not commissioned; externally peer reviewed.