


## A risk-stratified approach to colorectal cancer prevention and diagnosis

Mark A. Hull , Colin J. Rees, Linda Sharp and Sara Koo

**Abstract** | Population screening and endoscopic surveillance are used widely to prevent the development of and death from colorectal cancer (CRC). However, CRC remains a major cause of cancer mortality and the increasing burden of endoscopic investigations threatens to overwhelm some health services. This Perspective describes the rationale for and approach to improved risk stratification and decision-making for CRC prevention and diagnosis. Limitations of current approaches will be discussed using the UK as an example of the challenges faced by a particular health-care system, followed by discussion of novel risk biomarker utilization. We explore how risk stratification will be advantageous to current health-care providers and users, enabling more efficient use of limited colonoscopy resources. We discuss risk stratification in the setting of population screening as well as the surveillance of high-risk groups and investigation of symptomatic patients. We also address challenges in the development and validation of risk stratification tools and identify key research priorities.

Colorectal cancer (CRC) (or bowel cancer) is the second and fourth most common cancer diagnosis and cause of cancer death in Europe and the USA, respectively<sup>1–3</sup>. CRC mortality is 25–30% lower in locations where national bowel cancer screening programmes exist such as in the UK, Austria and Switzerland<sup>4</sup>. However, only a small proportion of CRCs are diagnosed through population-based screening programmes. For example, in the UK, only 10% of CRCs are diagnosed using population-based screening, with the vast majority diagnosed through symptomatic services (67%) or as an emergency presentation (23%)<sup>4</sup>.

Approximately 15 million colonoscopies were undertaken in the USA in 2012 and 900,000 were performed in the UK in 2016 (REFS<sup>5,6</sup>). The number of colonoscopies performed seems likely to continue to increase owing to the expansion of population screening programmes, a drive for earlier CRC diagnosis and an increasing surveillance burden from individuals identified as at risk of future colorectal neoplasia (on the basis of a genetic predisposition, for example, Lynch syndrome, or previous diagnosis of colorectal adenomas), which will

further stretch currently overburdened endoscopy services<sup>7,8</sup>. For example, in a UK national audit in 2011, more than 65% of colonoscopies were undertaken for diagnostic purposes, roughly 20% were undertaken for surveillance (ongoing assessment of a known high-risk group) and ~10% were undertaken during screening of average-risk members of the general population<sup>9</sup>. Overall, CRC was diagnosed in only 4% of procedures, one or more colorectal polyps were detected in 27% and 42% of examinations were normal<sup>9</sup>. There are several reasons for this low neoplasia yield: eligibility for ‘one size fits all’ population-based screening programmes is based solely on age; the evidence supporting the best surveillance interval for known high-risk groups (such as Lynch syndrome) is poor and is guided solely by colorectal adenoma characteristics<sup>10–12</sup>; and lower gastrointestinal symptoms correlate poorly with the presence of disease<sup>13,14</sup>. Guidelines such as those of the UK National Institute for Health and Care Excellence (NICE) on the investigation of suspected CRC are based on age, presence of specific symptoms and examination findings only, all of which are only weakly predictive of the presence of

CRC<sup>15</sup>. By comparison, Australian guidelines suggest triage of patients for colonoscopy based on age and symptoms but also consider the duration of symptoms and any faecal occult blood test result<sup>16</sup>.

The increasing demand in colonoscopy, driven by several factors, including the expansion of population-based screening, an ageing population and an extension of the indications for colonoscopy, coupled with restrictions to endoscopic services imposed by factors such as cost and the lengthy training required to achieve competence in high-quality colonoscopy, all of which have been exacerbated by the effects of coronavirus disease 2019 (COVID-19)<sup>17</sup>, mean that the need to avoid unnecessary investigations and treatment has never been more important. Thus, health-care services must further refine how patients are selected for colonoscopy beyond using existing (basic) criteria. A personalized medicine approach to disease treatment and prevention has been adopted in many areas of medicine; for example, the treatment of certain cancers based on their molecular tumour profile or the prevention of vascular events (myocardial infarction and stroke) based on lipid and blood pressure levels<sup>18,19</sup>. However, this approach has yet to be adopted for the prevention and diagnosis of CRC.

Systematic reviews of the several CRC risk prediction models developed to date for use in asymptomatic populations are available<sup>20,21</sup>. Thus, the main focus of this Perspective will be the latest developments towards a risk-stratified approach to CRC prevention and diagnosis, with particular emphasis on the emerging role of laboratory biomarkers for risk stratification. In general, we use the UK health-care system to exemplify current limitations in the prevention and diagnosis of CRC, but the concepts and developments that we discuss are broadly relevant and applicable to all advanced health-care systems worldwide.

### Benefits of risk stratification

The potential benefits of effective risk stratification for CRC prevention and early diagnosis are threefold: to individuals who are considered for investigation (henceforth, termed ‘patients’ for the purposes of this article), to health-care professionals and to service providers (FIG. 1).

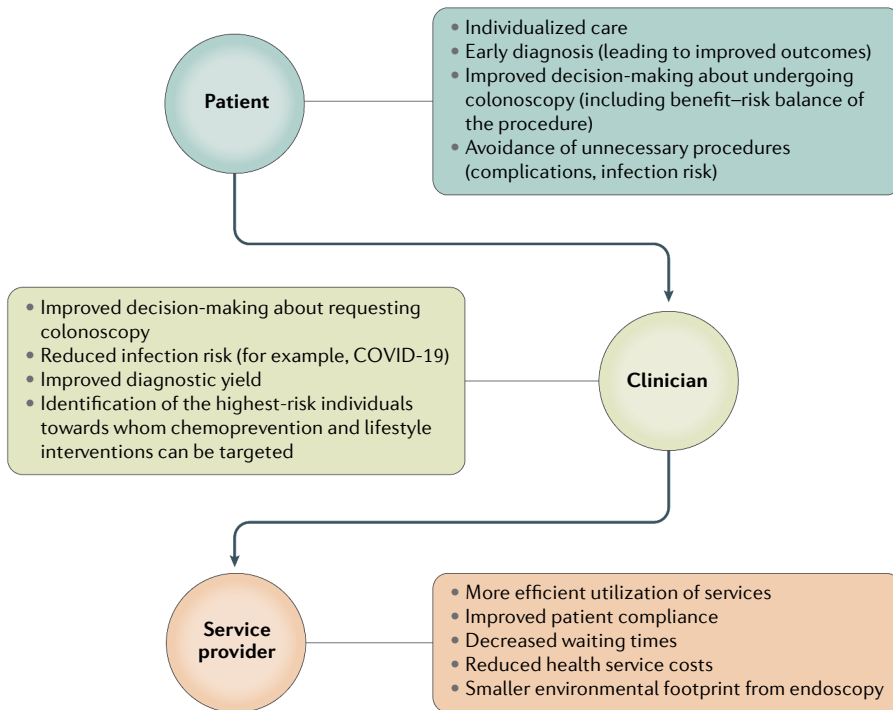


Fig. 1 | **Benefits of risk stratification for CRC prevention and diagnosis.** Overlapping benefits of improved risk stratification approaches for colorectal cancer (CRC) to patients, health-care professionals and service providers. For example, improved decision-making about investigation by patients and health-care professionals leads to better utilization of finite resources by institutions. COVID-19, coronavirus disease 2019.

Personal stratification for CRC risk has the potential to enable patients to better understand their individual risk of colorectal neoplasia and therefore to appreciate the need for colonoscopy, thereby improving patient compliance and attendance for colonoscopy. In a systematic review in 2015 of 11 randomized trials of 7 cancer risk tools (including decision aids for patients, self-completed risk assessment tools or a personalized risk prevention message) encompassing data on 7,677 patients and including 4 trials of CRC risk assessment in primary care, it was concluded that risk assessment tools were associated with improved patient knowledge, understanding of the importance of screening and the perception of CRC risk<sup>22</sup>. Additionally, in one of the studies, the intention to participate in CRC screening and screening uptake was increased in the intervention group compared with the control group (43% versus 5%)<sup>23</sup>. Advanced colorectal neoplasia is usually defined as a composite of CRC and ‘advanced’ colorectal adenomatous polyps (≥10 mm in size, or highly dysplastic lesions, or those with some villous histological architecture, considered clinically significant for detection based on the risk of malignant progression)<sup>24</sup>. The identification of those at highest risk

of colorectal neoplasia can translate into earlier diagnosis of advanced colorectal neoplasia, thereby improving patient outcomes based on the well-established association between earlier stage of CRC at diagnosis and better long-term outcomes, in addition to the importance of identifying and removing precancerous lesions<sup>24</sup>. Avoidance of colonoscopy in those at low risk of colorectal neoplasia will reduce exposure to an unpleasant procedure, which can be associated with substantial anxiety, pain or discomfort as well as carrying a risk (albeit small) of haemorrhage, perforation, other adverse events and even death<sup>25–27</sup>. With the advent of COVID-19, endoscopy avoidance also reduces infection risk in patients, many of whom are in high-risk groups for serious COVID-19 outcomes<sup>28,29</sup>.

For health-care professionals, risk stratification could help improve the identification of patients at greatest risk of harbouring or developing advanced colorectal neoplasia, thereby improving the diagnostic yield. This aspect has the further important benefit of identifying patients who might benefit from chemoprevention or other targeted prevention strategies such as lifestyle modification (for example, weight loss, dietary change or smoking cessation)<sup>20</sup>.

Risk stratification can enable tailoring of the post-polypectomy surveillance interval to the individual patient, which has been identified as an important area for further research<sup>12</sup>.

Lastly, for service providers, effective risk stratification should translate into improved efficiency and utilization of services, leading to a reduction in ‘wasted’ resources and thus allowing increased capacity for ‘appropriate’ procedures and decreased waiting times and health-care cost savings. Targeted colonoscopy should also help mitigate the substantial environmental footprint caused by endoscopy<sup>30</sup>.

Colonoscopy clearly has an important role in the diagnosis of many other conditions and in the management of non-cancer-related patient symptoms such as inflammatory bowel disease and investigation of diarrhoea<sup>31</sup>; however, these areas are beyond the scope of this article, which confines itself to the investigation of neoplasia.

**Current risk stratification**

A systematic review by Usher-Smith et al.<sup>20</sup> identified 52 CRC screening risk models encompassing 87 different variables. Peng et al.<sup>21</sup> focused their systematic review on 17 CRC risk models derived only from asymptomatic individuals undergoing colonoscopy. However, less than half of these (21 of 52 models) have been externally validated. Moreover, the discriminatory ability of models to identify individuals with colorectal neoplasia is variable<sup>20,21,32–34</sup>.

Risk prediction models in the context of CRC screening have used a median of five risk factors, with the most commonly used variables being age, sex, family history and BMI<sup>21</sup>. A few models have used genetic markers, including multiple SNPs identified by previous case-control studies<sup>20,21</sup>. In risk models in symptomatic patient cohorts, specific symptoms (rectal bleeding or altered bowel habit) or examination findings (abdominal mass) have also been included<sup>32</sup>. Although there is a known association between metabolic syndrome and increased CRC risk (relative risk ~1.4)<sup>35</sup>, there has been relatively little investigation of the potential value of clinical biomarkers of excess body weight and metabolic health (such as waist circumference, waist-hip ratio, diagnosis of diabetes and routinely available blood investigations such as levels of glycosylated haemoglobin and C-reactive protein) as predictors of risk. There has also been little investigation of whether or how ethnicity should be included in risk models, likely owing to the dearth of data from large cohorts with sufficient numbers

of different ethnicities included to enable interpretation<sup>20,21</sup>.

**Laboratory biomarkers.** A range of laboratory biomarkers have been or are currently being evaluated for use in CRC risk prediction models (TABLE 1). Faecal haemoglobin detection by the faecal immunochemical test (FIT) is the biomarker with most promise. It is already widely used in screening programmes internationally as a primary screening test or to triage individuals (subject to a positivity threshold) using symptomatic investigation algorithms, in which it can have a high specificity and negative predictive value for identifying CRC, depending on the threshold faecal haemoglobin value that is set<sup>36–38</sup>. For example, a 2019 systematic review of FIT use in asymptomatic individuals reported a sensitivity of 80% for CRC detection and a specificity of 91% at a positivity threshold of <10 µg/g (REF.<sup>38</sup>). In symptomatic patients, a primary care study in the UK reported a sensitivity and specificity of 91% for CRC for a FIT result ≥10 µg/g (REF.<sup>39</sup>). In this respect, FIT can be considered as the first quantitative laboratory biomarker in use for CRC risk stratification. The evaluation and use of FIT for risk stratification and triage for endoscopic investigation will undoubtedly accelerate as endoscopy utilization is reduced by COVID-19 control measures. The potential quantitative use of the absolute faecal haemoglobin level to predict CRC risk and all-cause mortality has been reported<sup>40–42</sup>. Other faecal biomarkers, such as faecal calprotectin levels, have been previously studied and have not been shown to be predictive of colorectal neoplasia in prospective studies<sup>43</sup>.

The multi-target faecal DNA test (based on the detection of mutant *KRAS*, methylated *BMP3* and methylation of the *NDRG4* promoter in faecal DNA) has excellent sensitivity for CRCs (>90%) but has limited performance in the detection of advanced colorectal adenomas (~40%) and is relatively expensive compared with faecal occult blood testing<sup>44</sup>. A clinical and economic comparison of the effectiveness of faecal DNA testing compared with FIT at a threshold haemoglobin value that provides similar diagnostic performance measures has yet to be performed<sup>45</sup>.

Intuitively, given the hereditary component of CRC risk, the incorporation of genetic markers should be associated with increased discriminatory power in risk prediction models. Genome-wide association studies have identified up to 70 independent

genetic loci associated with CRC risk<sup>46</sup>. A Genetic Risk Score (GRS) summarizes the effects of multiple SNPs associated with CRC and each GRS might vary depending on the number of SNPs included, the statistical analysis used and the model used to develop the risk score<sup>47</sup>. The addition of a GRS to CRC risk prediction models has been shown to improve the discriminatory power in some models addressing population risk<sup>48–50</sup>. For example, Iwasaki et al.<sup>49</sup> compared three models (amongst 349 men and 326 women with CRC compared with an equivalent number of age-matched controls without a CRC diagnosis): a non-genetic model (incorporating age, BMI, alcohol intake and smoking), a second model with age and GRS, and a third model incorporating the non-genetic model and GRS. The investigators found that the addition of GRS improved the predictive performance for CRC risk (c-statistic: 0.60, 0.63, 0.66, respectively) in men<sup>49</sup>. However, another study used data from the UK Biobank to construct a weighted GRS (which included 41 SNPs) and found that the inclusion of GRS to two previously developed risk models did not improve performance (c-statistic: Model 1, from 0.68 to 0.69; Model 2, from 0.67 to 0.67)<sup>51</sup>. A study of 95 common CRC genetic risk variants<sup>52</sup> concluded that the polygenic risk score predicted early-onset CRC, opening up the possibility

of risk stratification of young (<50 years) individuals for prevention measures.

The availability of technologies to assess the gut microbiota has led to the exploration of the predictive value of a faecal microbial signature to determine colorectal neoplastic risk. However, a review highlighted that studies to date are all small case-control or discovery cohort studies, with only a few studies including a validation cohort<sup>53</sup>. Despite these limitations, findings are promising; the diagnostic performance of tests based on both 16S ribosomal RNA and shotgun metagenomic sequencing as well as specific bacterial PCR panels is consistently good, with an area under the receiver-operator curve (AUROC) greater than 0.7 (REF.<sup>53</sup>). No study has yet to combine a specific microbiome profile or specific microbial signature with other risk factors (such as aspects of diet) in a multivariate risk prediction model, although a few studies have extended analyses to combine data with faecal occult blood testing<sup>53</sup>. The characterization of 'dysbiosis' associated with colorectal neoplasia and its link to dietary patterns is required before the incorporation of microbiome data into multivariate prediction models. The predictive value of the presence and abundance of certain pathobiont species (for example, *Fusobacterium nucleatum*) also remains to be determined in a clinical setting<sup>54,55</sup>.

Table 1 | Examples of potential laboratory biomarkers of CRC risk

Biomarker sample type	Examples	Sample collection	Other considerations
Blood	Circulating tumour DNA <sup>96</sup> ; genetic polymorphisms (SNPs) <sup>97</sup>	Requires venepuncture (clinic visit) but widely acceptable	Technically challenging and costly; polygenic SNPs relevant only for complex risk model
Faeces	Occult blood (gFOBT, FIT) <sup>62,63,65,66</sup> ; calprotectin <sup>43,98</sup> ; somatic DNA mutations <sup>99</sup> ; faecal microbiome <sup>100,101</sup> ; VOCs <sup>56</sup>	Noninvasive; not acceptable to some individuals; stool collection methodology critical for compliance	Occult blood testing and calprotectin assay widely available and already used in national screening programmes and in symptomatic services; high relative cost of DNA-based investigations
Urine	VOCs <sup>57,58</sup>	Noninvasive, easily obtainable	Unclear whether VOC signal relates to colorectal neoplasia and/or the background risk state of an individual (for example, metabolic health status)
Breath	VOCs <sup>102</sup>	Noninvasive, easily obtainable; suitable for point-of-care testing	Unclear whether VOC signal relates to colorectal neoplasia and/or the background risk state of an individual (for example, metabolic health status)

CRC, colorectal cancer; FIT, faecal immunochemical test; gFOBT, guaiac faecal occult blood test; VOCs, volatile organic compounds.

Volatile organic compounds (VOCs), which are gaseous substances released from biological samples such as urine, faeces and breath, are of interest for noninvasive cancer diagnosis, in which a profile of multiple VOCs could be associated with a particular cancer or pre-malignant state. A pilot study of 137 individuals found that a faecal VOC profile could be a potential biomarker for CRC<sup>56</sup>. Urinary VOCs have also been studied as an alternative noninvasive biomarker for CRC, although the evidence is limited at present<sup>57,58</sup>. One study has suggested that urinary VOCs could be used in combination with FIT, on the basis that it improves the diagnostic performance for CRC detection (sensitivity 0.97, specificity 0.77, negative predictive value (NPV) 1.0) compared with FIT alone (sensitivity 0.80, specificity 0.93, positive predictive value 0.44, NPV 0.9)<sup>56</sup>.

The complexity and size of datasets generated by ‘-omic’ analyses will be challenging for the incorporation of these laboratory biomarker outputs into clinically useful risk models. However, the rapid pace of progress in the fields of artificial intelligence and machine learning holds great promise for the future use of complex data alongside clinical factors in risk models<sup>59</sup>. Such tools could potentially enable multiple complex factors such as ‘-omics’ to be more rapidly assessed and clinically applied.

**For population CRC screening.** CRC screening programmes vary widely in different health-care systems<sup>60</sup>. For example, in the USA, Germany and Poland, screening is colonoscopy based but ‘opportunistic’ rather than population centred<sup>60</sup>. Guidelines by the European Society of Gastrointestinal Endoscopy recommend FIT-based screening<sup>61</sup>. Screening programmes are consistently focused on age<sup>60</sup>. However, rather than relying on age alone as the only risk factor for population screening for CRC, a more intelligent screening process could include risk factors such as sex, family history or BMI as well as other clinical biomarkers. For example, men have an increased incidence of CRC<sup>3</sup> and sex has consistently been included in symptomatic patient risk prediction models<sup>20,21</sup>. However, no factors are currently adopted into national screening programmes to modulate, for example, the frequency of screening or the FIT threshold value.

In the systematic review by Usher-Smith et al., the AUROC of studies using data from self-completed questionnaires, routine clinical data and/or genetic biomarkers were compared<sup>20</sup>. No clear trend was observed between an increasing number

of variables in a screening risk model and diagnostic performance<sup>20</sup>. The incorporation of laboratory blood results, such as haemoglobin levels, plasma glucose and lipid levels, or SNPs, is not consistently associated with an improved AUROC<sup>20</sup>. Encouragingly, two studies have found that the predictive ability of a positive threshold FIT in combination with other factors is better than that of a positive threshold FIT alone (Stegeman et al., AUROC increased from 0.69 to 0.76, variables were age, sex and history of previous screening; Cooper et al., AUROC increased from 0.63 to 0.66, variables were age, calcium intake, number of family members with CRC and smoking)<sup>62,63</sup>. With the increasing use of FIT in national CRC screening programmes, FIT-based risk stratification for population screening is an important area for immediate research, in particular exploring the utility of the absolute faecal haemoglobin concentration obtained from the FIT as a factor in multivariate risk stratification models<sup>64</sup>.

**For symptomatic patients.** Separate risk models have been developed for CRC risk prediction in symptomatic patients to aid diagnosis. These models have incorporated a limited number of patient-related factors (such as age, sex, family history or BMI) and usually include these factors in combination with symptoms (such as rectal bleeding or change in bowel habit), abnormal examination findings (such as rectal or abdominal mass) and routinely available blood test results<sup>65–70</sup>. In a systematic review, risk models in symptomatic patients had a good discriminatory ability for the presence of CRC (AUROC 0.83–0.97)<sup>32</sup>.

Symptoms alone are poorly discriminatory for CRC, primarily because lower gastrointestinal symptoms are so ubiquitous. However, rectal bleeding and weight loss have a stronger association with CRC than other non-specific symptoms such as change in bowel habit or abdominal pain<sup>71</sup>. The addition of the presence or absence of these symptoms into a risk model containing age, sex and medical history has been associated with an increased diagnostic performance (AUROC increased from 0.79 to 0.85)<sup>70</sup>.

The use and evaluation of FIT in the symptomatic population is increasing worldwide, with intense interest driven by the reduced endoscopic provision forced by infection control measures due to COVID-19. Current evidence suggests that a low threshold FIT (<12 µg/g) has a

high NPV for CRC (99%)<sup>43,69</sup> and adenomas (94%)<sup>43</sup>. However, it has lower sensitivity and specificity for adenomas (69% and 56%, respectively) than for CRC (84% and 93%, respectively), suggesting that FIT is best placed to exclude clinically significant (advanced) colorectal neoplasia in symptomatic patients to avoid unnecessary investigations<sup>43</sup>.

Risk prediction models, such as the Bristol–Birmingham equation, the COLONPREDICT model or the FAST model, are superior at detecting CRC than the NICE referral guidelines: AUROC for the Bristol–Birmingham equation was 0.83 (0.82–0.84) versus 0.64 for NICE guideline CG27 (REF.<sup>68</sup>); AUROC for COLONPREDICT was 0.92 (0.91–0.94) versus 0.53 for NICE guideline NG12 (REF.<sup>65</sup>); and AUROC for FAST was 0.87 versus 0.53 for NICE guideline NG12 (REF.<sup>65</sup>). Interestingly, FIT level alone was also superior to NICE guideline NG12 (AUROC 0.86 and 0.53, respectively)<sup>65</sup>.

**For CRC surveillance procedures.**

Surveillance colonoscopy places considerable demands upon endoscopic services. Indeed, the need to rationalize surveillance colonoscopy for CRC risk is recognized internationally and is deemed a key research priority<sup>12,72–74</sup>. However, even under the new UK guidelines, surveillance will still be determined entirely by polyp number and characteristics, which have been shown to have limited predictive value for subsequent CRC risk<sup>12,75</sup>. Improved risk stratification using patient factors has the potential to enable clinicians to identify a genuine high-risk population for careful surveillance and to avoid procedures in lower-risk patients who might require less frequent surveillance or a non-colonoscopy approach to surveillance<sup>76</sup>.

High-risk patients could particularly benefit from healthy lifestyle advice and support to modify behaviour known to prevent the occurrence or recurrence of neoplasia<sup>77</sup>. Research into the effects of modifiable lifestyle risk factor interventions linked to screening and surveillance on colorectal neoplastic risk is needed. Additionally, chemoprevention could be targeted at this group of patients (discussed later).

The role of FIT to risk-stratify for surveillance interval has been evaluated in a small number of studies. Terhaar sive Droste et al. evaluated FIT at a threshold of 50 ng/ml haemoglobin levels in patients attending surveillance colonoscopy because of a previous history of adenomas or



CRC and a family history of CRC and found that the sensitivity was 80% for CRC but that the sensitivity for advanced adenomas was poor, at 28%<sup>78</sup>. Additionally, a second FIT sample prior to colonoscopy only slightly improved the sensitivity for advanced adenoma (33%)<sup>78</sup>. However, the NPVs for CRC and colorectal adenomas were high (99.9% and 92%, respectively)<sup>78</sup>. Similarly, the 'FIT for follow-up' study by Cross et al. predicted that, if annual FIT replaced three-yearly colonoscopic surveillance in patients with previous multiple colorectal adenomas, a substantial proportion of advanced neoplasia would be missed (28%)<sup>76</sup>.

To date, no risk model to direct personalized surveillance has been described, making this a priority area for further research.

#### **Risk-based chemoprevention.**

Chemoprevention (the use of drugs or natural substances to prevent cancer) is a recognized prevention strategy for CRC<sup>79</sup>. Despite compelling evidence for the anti-CRC activity of aspirin, it is currently not used for primary or secondary prevention of CRC due to the rare but clinically significant risk of bleeding, both gastrointestinal and cerebral<sup>80</sup>. Furthermore, clinicians are currently unable to balance individual risk and benefit from aspirin due to the lack of a predictive model for aspirin chemoprevention efficacy. Additionally, there is a lack of understanding for how best to utilize chemoprevention in an effective way in combination with endoscopic screening and surveillance<sup>81</sup>. To date, the US Preventive Services Task Force has recommended the use of aspirin for CRC prevention in individuals with an existing increased risk of cardiovascular events and NICE recommended aspirin in individuals with Lynch syndrome in the UK<sup>82,83</sup>. The demonstration that aspirin and the omega-3 fatty acid eicosapentaenoic acid have differential preventive activity against conventional colorectal adenomas compared with serrated polyps in a high-risk patient group<sup>84</sup>, along with data suggesting that aspirin dose should be tailored to body weight<sup>85</sup>, should stimulate a precision medicine approach to chemoprevention. This approach will require the development of a risk tool that will need to balance potential harm against cancer and the overall mortality risk reduction.

Other CRC chemoprevention agents, including the non-steroidal anti-inflammatory drugs sulindac and celecoxib, have been used in patients with

the rare CRC predisposition syndrome familial adenomatous polyposis based on evidence from polyp prevention trials<sup>86</sup>. However, the effectiveness of these and other chemoprevention agents for the delay or avoidance of colectomy as well as for the reduction in CRC risk itself has not been evaluated to date.

#### **An ideal CRC risk model**

**Development and validation.** An ideal model would constitute only factors that are strongly associated (either positively or negatively) with CRC risk (and are independent of, or at least not highly correlated with, one another) so that the model can discriminate between those with and without disease. In reality, things are not this simple. The value of a predictor depends not only on the strength of its association with the outcome but also on its distribution and/or frequency in the population (which is one reason why models developed in different populations vary). In addition, many clinical CRC risk factors are related; for example, excess body weight and diabetes<sup>87</sup>.

During the assessment of risk model performance, internal and external validation should be undertaken. Internal validation assesses the model utilizing the data upon which it was derived through methods such as bootstrapping or cross-validation. External validation assesses the model using a new independent dataset from a different cohort of patients<sup>88</sup>. However, many CRC models lack external validation. Given the existence of different CRC risk prediction models, a direct comparison of the performance of different models in the same validation cohort is also needed<sup>89</sup>. Lastly, model development should be reported in a clear and transparent way, adhering to the TRIPOD statement and checklist for the reporting of multivariable prediction model studies<sup>90</sup>.

**Implementation and barriers.** Risk stratification is likely to complement and inform clinical practice rather than replace clinical decision-making. An ideal model for use in practice would have face validity, be quick and easy to use, fit seamlessly within standard consultations, and generate easily understood results, enabling health-care professionals to categorize patients into different risk groups. In addition, obtaining the data required to estimate a patient's risk should not involve additional health-care episodes, expensive tests or tests that are poorly tolerated by patients.

Thus far, although some published risk models perform at a level that is considered 'clinically acceptable' (AUROC >0.7), these models have yet to be adopted in clinical practice. One probable explanation for the lack of clinical implementation is the absence of evidence that the use of a CRC risk prediction model in clinical practice will enhance clinical decision-making and improve health outcomes for patients. Kappen et al. argued that prospective comparative studies (ideally cluster randomized trials) are needed to quantify the impact on clinical decision-making of risk prediction models and suggested various considerations in the conduct of these so-called impact studies<sup>91</sup>.

Further reasons for the non-adoption of CRC risk prediction models probably include the lack of knowledge about: acceptability to patients, health professionals in primary and secondary care, and service providers; feasibility, especially in terms of obtaining the information on the risk factors that comprise the models, and use within routine consultations or screening programmes; and cost-effectiveness compared with standard of care. One study that tested the QCancer Risk tool in a primary care setting in Australia identified a range of practical barriers around implementation in daily practice, including difficulties in introducing the tool into the workflow of the consultation<sup>92</sup>. In addition, some users found that the output of the risk model was 'confrontational', confusing and alarming to patients and others (often the more experienced practitioners) distrusted the model output when it conflicted with their clinical judgment<sup>92</sup>. Similar findings emerged from another study by the same group that explored issues around the implementation of the CRISP CRC risk prediction tool<sup>93</sup>. This finding suggests that considerable work is still needed to address the barriers to implementation before these models can be widely adopted.

A concern for any risk prediction model is the scenario in which a patient is categorized as low-risk but harbours clinically significant, advanced colorectal neoplasia or, conversely, is categorized as high-risk but does not have neoplasia. Clearly 'low risk' does not equate to 'no risk' and 'high risk' does not equate to cancer. All tests or scores (even the most well accepted and commonly used biochemical markers) can result in a false-positive or false-negative result. For example, in the study by Stegeman et al. of 101 patients with advanced neoplasia, 7% of FIT-negative individuals were 'risk positive' according

Box 1 | Research priorities for improved risk stratification for CRC

- The clinical research community needs to perform external validation of colorectal cancer (CRC) risk prediction tools and facilitate the direct comparison of different risk prediction models.
- There is a need to evaluate the influence of risk prediction models on clinical decision-making and patient outcomes (including assessment of cost-effectiveness and any potential harms) in multiple settings, namely population screening, surveillance of high-risk groups and early, efficient diagnosis of CRC in symptomatic individuals.
- The acceptability and feasibility of a risk-stratified approach to CRC prevention to patients, clinicians and health-care providers need investigation.
- Risk prediction models of asymptomatic and symptomatic patients require modelling around the delivery of symptomatic and screening services to ensure optimal service delivery of screening and symptomatic care.
- There is a need for artificial intelligence and machine learning development to link large individual datasets (for example, faecal microbiome profile, volatile organic compounds) to derive clinically useful prediction models.

to the model (FIT plus risk questionnaire), whereas 2% were FIT positive but ‘risk negative’<sup>63</sup>. These issues need to be appreciated by those using the models in practice and conveyed to patients in a way that is understandable and minimizes both the possibility of false reassurance or cancer worry.

A further potential concern is that providing patients with personalized cancer risk information could cause psychological harm. The limited work on this topic (most of which is not focused on CRC) suggests that using self-completed risk assessment tools or personalized risk prevention messages might not result in increased patient anxiety prior to the investigation<sup>22,94</sup>; however, further research specific to CRC diagnosis and prevention is needed.

Finally, a challenge remains with regards to dealing with conflict between the results of current widely accepted testing (for example, FIT screening results) and results from a risk-prediction model. For example, strategies for managing those with high-risk scores but negative or low FIT results and those with low-risk scores and positive or high FIT results are needed as both scenarios conflict with current routine practice and could have the potential to lead to harms for the patient (such as unnecessary procedures or missed cancer).

**Challenges and research priorities**

Despite the multiple potential benefits of risk stratification in CRC, as yet, there is insufficient evidence demonstrating the clinical effectiveness and cost-effectiveness of risk stratification in different settings, including screening, surveillance and assessment of symptomatic patients, compared with current practice. It is unlikely that a ‘one size fits all’ model will be developed that is applicable in multiple settings and health-care systems.

We have identified several priorities for research that should hasten the use of a risk-stratified approach to the prevention of CRC (BOX 1). The development of acceptable and feasible risk-stratified approaches to CRC prevention, screening and early diagnosis need to be prioritized. Large, prospective cohort data or, ideally, anonymized data from national databases of sufficient quality (since these would be population based and not subject to selection biases) are required to provide sufficient power to evaluate individual factors in multivariate models, to undertake external validation of developed models and to compare the performance of different models. There could be a role for artificial intelligence in analysis as demonstrated by Nartowt et al., who showed the potential of this approach for CRC based on information available on medical databases (such as age, sex, family history of CRC or presence of ischaemic heart disease)<sup>95</sup>.

The research community needs to embrace the testing effectiveness of ‘real-world’ screening, surveillance and early diagnosis risk stratification algorithms in terms of their influence on clinical decision-making and patient outcomes. The effects on health-care services and patient care as well as on patient experience of care need careful evaluation. However, we also need to better understand how patients and clinicians perceive the utilization of individual risk stratification (with the attendant risk of missed lesions or psychological harms of being labelled high-risk) compared with the traditional approach of considering population risk for health-care provision. It is important that research into improved CRC risk stratification is not limited to endoscopic prevention measures but that it also addresses other prevention interventions, including better, more targeted use of

chemoprevention agents and lifestyle modification. Ultimately, risk stratification has the potential to modify the patient journey by the individualization of patient care. Further work remains to identify clinical markers and new biomarkers that would provide the best yield of neoplasia for patients attending screening, surveillance and diagnostic colonoscopies.

**Conclusions**

Improved risk stratification for the prevention and early diagnosis of CRC is essential to best utilize stretched clinical services and improve outcomes. This approach will enable resources to be targeted optimally and reduce unnecessary investigations in low-risk individuals. The potential of combining knowledge of clinical risk factors for CRC with laboratory biomarkers, such as FIT and VOC profiles, is only just now being realized. Careful reporting and validation of multivariable risk prediction models in individual clinical settings will be essential in order to derive clinically relevant tools for the precision diagnosis and prevention of this common malignancy.

Mark A. Hull<sup>1</sup>✉, Colin J. Rees<sup>2</sup>, Linda Sharp<sup>2</sup> and Sara Koo<sup>2</sup>

<sup>1</sup>Leeds Institute of Medical Research at St James’s, University of Leeds, Leeds, UK.

<sup>2</sup>Population Health Sciences Institute, Newcastle University, Newcastle, UK.

✉e-mail: m.a.hull@leeds.ac.uk

<https://doi.org/10.1038/s41575-020-00368-3>

Published online 16 October 2020

1. Ait Ouakrim, D. et al. Trends in colorectal cancer mortality in Europe: retrospective analysis of the WHO mortality database. *BMJ* **351**, h4970 (2015).
2. Centers for Disease Control and Prevention. United States Cancer Statistics: Data Visualizations <https://gis.cdc.gov/Cancer/USCS/DataViz.html> (2015).
3. Cancer Research UK. Bowel Cancer Statistics <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-One> (2017).
4. Public Health England: National Cancer Intelligence Network. Routes to diagnosis 2006–2016 year breakdown [http://www.ncin.org.uk/publications/routes\\_to\\_diagnosis](http://www.ncin.org.uk/publications/routes_to_diagnosis) (2018).
5. Shenbagaraj, L. et al. Endoscopy in 2017: a national survey of practice in the UK. *Frontline Gastroenterol.* **10**, 7–15 (2019).
6. Joseph, D. A. et al. Colorectal cancer screening: estimated future colonoscopy need and current volume and capacity. *Cancer* **122**, 2479–2486 (2016).
7. Kaur A. Bowel Cancer UK. Diagnosing bowel cancer early – a service at breaking point <https://www.bowelcanceruk.org.uk/news-and-blogs/campaigns-and-policy-blog/diagnosing-bowel-cancer-early-a-service-at-breaking-point/> (2019).
8. Centre for Workforce Intelligence. *Securing the Future Workforce Supply: Gastrointestinal Endoscopy Workforce Review* (CFWI, 2017).
9. Gavin, D. R. et al. The national colonoscopy audit: a nationwide assessment of the quality and safety of colonoscopy in the UK. *Gut* **62**, 242–249 (2013).
10. Hassan, C. et al. Post-polypectomy colonoscopy surveillance: European society of gastrointestinal endoscopy (ESGE) guideline. *Endoscopy* **45**, 842–851 (2013).

11. Lieberman, D. A. et al. Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US multi-society task force on colorectal cancer. *Gastroenterology* **143**, 844–857 (2012).
12. Rutter, M. D. et al. British Society of Gastroenterology/ Association of Coloproctology of Great Britain and Ireland/Public Health England post-polypectomy and post-colorectal cancer resection surveillance guidelines. *Gut* **69**, 201–223 (2020).
13. Vulliamy, P., McCluney, S., Raouf, S. & Banerjee, S. Trends in urgent referrals for suspected colorectal cancer: an increase in quantity, but not in quality. *Ann. R. Coll. Surg. Engl.* **98**, 564–567 (2016).
14. Ford, A. C. et al. Diagnostic utility of alarm features for colorectal cancer: systematic review and meta-analysis. *Gut* **57**, 1545–1552 (2008).
15. National Institute of Health and Care Excellence. *Suspected Cancer: Recognition and Referral. NICE Guidance (NG12)* (NICE, 2017).
16. Cancer Council Australia. Clinical practice guidelines for the prevention, early detection and management of colorectal cancer. [https://wiki.cancer.org.au/australia/Guidelines:Colorectal\\_cancer](https://wiki.cancer.org.au/australia/Guidelines:Colorectal_cancer) (2018).
17. Han, J. et al. Preventing the spread of COVID-19 in digestive endoscopy during the resuming period: meticulous execution of screening procedures. *Gastrointest. Endosc.* <https://doi.org/10.1016/j.gie.2020.03.3855> (2020).
18. Pernas, S. & Tolaney, S. M. HER2-positive breast cancer: new therapeutic frontiers and overcoming resistance. *Ther. Adv. Med. Oncol.* **11**, 1758835919833519 (2019).
19. Hippisley-Cox, J. et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* **336**, 1475–1482 (2008).
20. Usher-Smith, J. A., Walter, F. M., Emery, J. D., Win, A. K. & Griffin, S. J. Risk prediction models for colorectal cancer: a systematic review. *Cancer Prev. Res.* **9**, 13–26 (2016).
21. Peng, L., Weigl, K., Boakye, D. & Brenner, H. Risk scores for predicting advanced colorectal neoplasia in the average-risk population: a systematic review and meta-analysis. *Am. J. Gastroenterol.* **113**, 1788–1800 (2018).
22. Walker, J. G., Licqurish, S., Chiang, P. P. C., Pirotta, M. & Emery, J. D. Cancer risk assessment tools in primary care: a systematic review of randomized controlled trials. *Ann. Fam. Med.* **13**, 480–489 (2015).
23. Schroy, P. C. et al. Aid-assisted decision making and colorectal cancer screening: a randomized controlled trial. *Am. J. Prev. Med.* **43**, 573–583 (2012).
24. Gill, M. D. et al. Comparison of screen-detected and interval colorectal cancers in the Bowel Cancer Screening Programme. *Br. J. Cancer* **107**, 417–421 (2012).
25. Neilson, L. et al. Patient experience of gastrointestinal endoscopy: informing the development of the Newcastle ENDOPREM™. *Frontline Gastroenterol.* **11**, 209–217 (2020).
26. Warren, J. L. et al. Adverse events after outpatient colonoscopy in the Medicare population. *Ann. Intern. Med.* **150**, 849–857 (2009).
27. Gatto, N. M. et al. Risk of perforation after colonoscopy and sigmoidoscopy: a population-based study. *J. Natl Cancer Inst.* **95**, 230–236 (2003).
28. Chiu, P. W. Y. et al. Practice of endoscopy during COVID-19 pandemic: position statements of the Asian Pacific society for digestive endoscopy (APSDE-COVID statements). *Gut* <https://doi.org/10.1136/gutjnl-2020-321185> (2020).
29. Gralnek, I. M. et al. ESGE and ESGENA position statement on gastrointestinal endoscopy and the COVID-19 pandemic. *Endoscopy* <https://doi.org/10.1055/a-1155-6229> (2020).
30. Maurice, J. et al. Green endoscopy: a call for sustainability in the midst of COVID-19. *Lancet Gastroenterol. Hepatol.* **5**, 636–638 (2020).
31. Bowles, C. J. A. et al. A prospective study of colonoscopy practice in the UK today: are we adequately prepared for national colorectal cancer screening tomorrow? *Gut* **53**, 277–283 (2004).
32. Williams, T. G. S., Cubiella, J., Griffin, S. J., Walter, F. M. & Usher-Smith, J. A. Risk prediction models for colorectal cancer in people with symptoms: a systematic review. *BMC Gastroenterol.* **16**, 63 (2016).
33. Ma, G. K. & Ladabaum, U. Personalizing colorectal cancer screening: a systematic review of models to predict risk of colorectal neoplasia. *Clin. Gastroenterol. Hepatol.* **12**, 1624–1634 (2014).
34. Smith, T. et al. Comparison of prognostic models to predict the occurrence of colorectal cancer in asymptomatic individuals: a systematic literature review and external validation in the EPIC and UK Biobank prospective cohort studies. *Gut* **68**, 672–683 (2019).
35. Esposito, K. et al. Colorectal cancer association with metabolic syndrome and its components: a systematic review with meta-analysis. *Endocrine* **44**, 634–647 (2013).
36. Lee, J. K., Liles, E. G., Bent, S., Levin, T. R. & Corley, D. A. Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis. *Ann. Intern. Med.* **160**, 171 (2014).
37. Westwood, M. et al. Faecal immunochemical tests (FIT) can help to rule out colorectal cancer in patients presenting in primary care with lower abdominal symptoms: a systematic review conducted to inform new NICE DG30 diagnostic guidance. *BMC Med.* **15**, 189 (2017).
38. Selby, K. et al. Effect of sex, age, and positivity threshold on fecal immunochemical test accuracy: a systematic review and meta-analysis. *Gastroenterology* **157**, 1494–1505 (2019).
39. Nicholson, B. D. et al. Faecal immunochemical testing for adults with symptoms of colorectal cancer attending English primary care: a retrospective cohort study of 14487 consecutive test requests. *Aliment. Pharmacol. Ther.* <https://doi.org/10.1111/apt.15969> (2020).
40. Senore, C. et al. Performance of colorectal cancer screening in the European Union Member States: data from the second European screening report. *Gut* **68**, 1232–1244 (2019).
41. Selby, K. et al. Influence of varying quantitative fecal immunochemical test positivity thresholds on colorectal cancer detection: a community-based cohort study. *Ann. Intern. Med.* **169**, 439–447 (2018).
42. Libby, G. et al. Occult blood in faeces is associated with all-cause and non-colorectal cancer mortality. *Gut* **67**, 2116–2123 (2018).
43. Widlak, M. M. et al. Diagnostic accuracy of faecal biomarkers in detecting colorectal cancer and adenoma in symptomatic patients. *Aliment. Pharmacol. Ther.* **45**, 354–363 (2017).
44. Imperiale, T. F. et al. Multitarget stool DNA testing for colorectal-cancer screening. *N. Engl. J. Med.* **370**, 1287–1297 (2014).
45. Senore, C. & Segnan, N. Multitarget stool DNA testing for colorectal-cancer screening. *N. Engl. J. Med.* **371**, 184–188 (2014).
46. Law, P. J. et al. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat. Commun.* **10**, 2154 (2019).
47. Sugrue, L. P. & Desikan, R. S. What are polygenic scores and why are they important? *JAMA* **321**, 1820–1821 (2019).
48. Xin, J. et al. Evaluating the effect of multiple genetic risk score models on colorectal cancer risk prediction. *Gene* **673**, 174–180 (2018).
49. Iwasaki, M. et al. Inclusion of a genetic risk score into a validated risk prediction model for colorectal cancer in Japanese men improves performance. *Cancer Prev. Res.* **10**, 535–541 (2017).
50. Weigl, K. et al. Strongly enhanced colorectal cancer risk stratification by combining family history and genetic risk score. *Clin. Epidemiol.* **10**, 143–152 (2018).
51. Smith, T., Gunter, M. J., Tzoulaki, I. & Muller, D. C. The added value of genetic information in colorectal cancer risk prediction models: development and evaluation in the UK Biobank prospective cohort study. *Br. J. Cancer* **119**, 1036–1039 (2018).
52. Archambault, A. N. et al. Cumulative burden of colorectal cancer-associated genetic variants is more strongly associated with early-onset vs late-onset cancer. *Gastroenterology* **158**, 1274–1286 (2020).
53. Wong, S. H. & Yu, J. Gut microbiota in colorectal cancer: mechanisms of action and clinical applications. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 690–704 (2019).
54. Idrissi Janati, A., Karp, I., Sabri, H. & Emami, E. Is a fusobacterium nucleatum infection in the colon a risk factor for colorectal cancer?: a systematic review and meta-analysis protocol. *Syst. Rev.* **8**, 114 (2019).
55. Shang, F.-M. & Liu, H.-L. Fusobacterium nucleatum and colorectal cancer: a review. *World J. Gastrointest. Oncol.* **10**, 71–81 (2018).
56. Bond, A. et al. Volatile organic compounds emitted from faeces as a biomarker for colorectal cancer. *Aliment. Pharmacol. Ther.* **49**, 1005–1012 (2019).
57. Widlak, M. M. et al. Risk stratification of symptomatic patients suspected of colorectal cancer using faecal and urinary markers. *Colorectal Dis.* **20**, O335–O342 (2018).
58. Arasaradnam, R. P. et al. Detection of colorectal cancer (CRC) by urinary volatile organic compound analysis. *PLoS ONE* **9**, e108750 (2014).
59. van der Sommen, F. et al. Machine learning in GI endoscopy: practical guidance in how to interpret a novel field. *Gut* <https://doi.org/10.1136/gutjnl-2019-320466> (2020).
60. Schreuders, E. H. et al. Colorectal cancer screening: a global overview of existing programmes. *Gut* **64**, 1637–1649 (2015).
61. Saftoiu, A. et al. Role of gastrointestinal endoscopy in the screening of digestive tract cancers in Europe: European Society of Gastrointestinal Endoscopy (ESGE) position statement. *Endoscopy* **52**, 293–304 (2020).
62. Cooper, J. A. et al. Risk-adjusted colorectal cancer screening using the FIT and routine screening data: development of a risk prediction model. *Br. J. Cancer* **118**, 285–293 (2018).
63. Stegeman, I. et al. Combining risk factors with faecal immunochemical test outcome for selecting CRC screenees for colonoscopy. *Gut* **63**, 466–471 (2014).
64. van de Veerdonk, W., Hoek, S., Peeters, M. & Van Hal, G. Towards risk-stratified colorectal cancer screening. Adding risk factors to the fecal immunochemical test: evidence, evolution and expectations. *Prev. Med.* **126**, 105746 (2019).
65. Herrero, J.-M., Vega, P., Salve, M., Bujanda, L. & Cubiella, J. Symptom or faecal immunochemical test based referral criteria for colorectal cancer detection in symptomatic patients: a diagnostic tests study. *BMC Gastroenterol.* **18**, 155 (2018).
66. Cubiella, J. et al. Development and external validation of a faecal immunochemical test-based prediction model for colorectal cancer detection in symptomatic patients. *BMC Med.* **14**, 128 (2016).
67. Cubiella, J. et al. The fecal hemoglobin concentration, age, and sex test score: development and external validation of a simple prediction tool for colorectal cancer detection in symptomatic patients. *Int. J. Cancer* **140**, 2201–2211 (2017).
68. Marshall, T. et al. The diagnostic performance of scoring systems to identify symptomatic colorectal cancer compared to current referral guidance. *Gut* **60**, 1242–1248 (2011).
69. Hippisley-Cox, J. & Coupland, C. Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. *Br. J. Gen. Pract.* **62**, e29–e37 (2012).
70. Adelstein, B.-A. et al. Who needs colonoscopy to identify colorectal cancer? Bowel symptoms do not add substantially to age and other medical history. *Aliment. Pharmacol. Ther.* **32**, 270–281 (2010).
71. Adelstein, B.-A., Macaskill, P., Chan, S. F., Katelaris, P. H. & Irwig, L. Most bowel cancer symptoms do not indicate colorectal cancer and polyps: a systematic review. *BMC Gastroenterol.* **11**, 65 (2011).
72. Ladabaum, U. & Schoen, R. E. Post-polypectomy surveillance that would please goldilocks—not too much, not too little, but just right. *Gastroenterology* **150**, 791–796 (2016).
73. Winawer, S. J. & Zauber, A. G. Can post-polypectomy surveillance be less intensive? *Lancet Oncol.* **18**, 707–709 (2017).
74. Rees, C. J. et al. European Society of Gastrointestinal Endoscopy — establishing the key unanswered research questions within gastrointestinal endoscopy. *Endoscopy* **48**, 884–891 (2016).
75. Cross, A. J. et al. Long-term colorectal cancer incidence after adenoma removal and the effects of surveillance on incidence: a multicentre, retrospective, cohort study. *Gut* **69**, 1645–1658 (2020).
76. Cross, A. J. et al. Faecal immunochemical tests (FIT) versus colonoscopy for surveillance after screening and polypectomy: a diagnostic accuracy and cost-effectiveness study. *Gut* **68**, 1642–1652 (2019).
77. Song, M., Chan, A. T. & Sun, J. Influence of gut microbiome, diet, and environment on risk of colorectal cancer. *Gastroenterology* **158**, 322–340 (2020).
78. Terhaar sive Droste, J. et al. Faecal immunochemical test accuracy in patients referred for surveillance colonoscopy: a multi-centre cohort study. *BMC Gastroenterol.* **12**, 94 (2012).
79. Keum, N. N. & Giovannucci, E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 713–732 (2019).
80. Cuzick, K. et al. Estimates of benefits and harms of prophylactic use of aspirin in the general population. *Ann. Oncol.* **26**, 47–57 (2015).
81. Drew, D. A., Cao, Y. & Chan, A. T. Aspirin and colorectal cancer: the promise of precision chemoprevention. *Nat. Rev. Cancer* **16**, 173–189 (2016).

82. National Institute of Health and Care Excellence. Offer daily aspirin to those with inherited genetic condition to reduce the risk of colorectal cancer <https://www.nice.org.uk/news/article/offer-daily-aspirin-to-those-with-inherited-genetic-condition-to-reduce-the-risk-of-colorectal-cancer> (2019).
83. US Preventative Services Task Force. Aspirin use to prevent cardiovascular disease and colorectal cancer: preventive medication <https://www.uspreventiveservicestaskforce.org/Page/Document/UpdateSummaryFinal/aspirin-to-prevent-cardiovascular-disease-and-cancer> (2016).
84. Hull, M. A. et al. Eicosapentaenoic acid and aspirin, alone and in combination, for the prevention of colorectal adenomas (seAFOod Polyp Prevention trial): a multicentre, randomised, double-blind, placebo-controlled, 2x2 factorial trial. *Lancet* **392**, 2583–2594 (2018).
85. Rothwell, P. M. et al. Effects of aspirin on risks of vascular events and cancer according to bodyweight and dose: analysis of individual patient data from randomised trials. *Lancet* **392**, 387–399 (2018).
86. Ricciardiello, L., Ahnen, D. J. & Lynch, P. M. Chemoprevention of hereditary colon cancers: time for new strategies. *Nat. Rev. Gastroenterol. Hepatol.* **13**, 352–361 (2016).
87. Lega, I. C. & Lipscombe, L. L. Review: diabetes, obesity and cancer- pathophysiology and clinical implications. *Endocr. Rev.* **41**, 33–52 (2020).
88. Grant, S. W., Collins, G. S. & Nashef, S. A. M. Statistical primer: developing and validating a risk prediction model. *Eur. J. Cardiothorac. Surg.* **54**, 203–208 (2018).
89. Collins, G. S. & Moons, K. G. M. Comparing risk prediction models. *BMJ* **344**, e3186 (2012).
90. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. M. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ* **350**, g7594 (2015).
91. Kappen, T. H. et al. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagnostic Progn. Res.* **2**, 11 (2018).
92. Chiang, P. P.-C., Giance, D., Walker, J., Walter, F. M. & Emery, J. D. Implementing a Qcancer risk tool into general practice consultations: an exploratory study using simulated consultations with Australian general practitioners. *Br. J. Cancer* **112**, S77–S83 (2015).
93. Walker, J. G. et al. The CRISP colorectal cancer risk prediction tool: an exploratory study using simulated consultations in Australian primary care. *BMC Med. Inf. Decis. Mak.* **17**, 13 (2017).
94. French, D. P. et al. Psychological impact of providing women with personalised 10-year breast cancer risk estimates. *Br. J. Cancer* **118**, 1648–1657 (2018).
95. Nartowt, B. J. et al. Scoring colorectal cancer risk with an artificial neural network based on self-reportable personal health data. *PLoS ONE* **14**, e0221421 (2019).
96. Bach, S. et al. Circulating tumor DNA analysis: clinical implications for colorectal cancer patients. A systematic review. *JNCI Cancer Spectr.* <https://doi.org/10.1093/jncics/pkz042> (2019).
97. Wen, J., Xu, Q. & Yuan, Y. Single nucleotide polymorphisms and sporadic colorectal cancer susceptibility: a field synopsis and meta-analysis. *Cancer Cell Int.* **18**, 155 (2018).
98. Turvill, J. et al. Diagnostic accuracy of one or two faecal haemoglobin and calprotectin measurements in patients with suspected colorectal cancer. *Scand. J. Gastroenterol.* **53**, 1526–1534 (2018).
99. Lin, S.-H. et al. The somatic mutation landscape of premalignant colorectal adenoma. *Gut* **67**, 1299–1305 (2018).
100. Saus, E. et al. Microbiome and colorectal cancer: roles in carcinogenesis and clinical potential. *Mol. Asp. Med.* **69**, 93–106 (2019).
101. Dahmus, J. D., Kotler, D. L., Kastenber, D. M. & Kistler, C. A. The gut microbiome and colorectal cancer: a review of bacterial pathogenesis. *J. Gastrointest. Oncol.* **9**, 769–777 (2018).
102. Markar, S. R. et al. Breath volatile organic compound profiling of colorectal cancer using selected ion flow-tube mass spectrometry. *Ann. Surg.* **269**, 903–910 (2019).

**Author contributions**

The authors contributed equally to all aspects of the article.

**Competing interests**

The authors declare no competing interests.

**Peer review information**

*Nature Reviews Gastroenterology & Hepatology* thanks H. Brenner, T. Levin, R. Steele and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Publisher's note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2020