

Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia

Orouba Almilaji,^{1,2} Carla Smith,¹ Sue Surgenor,¹ Andrew Clegg,³ Elizabeth Williams,¹ Peter Thomas,² Jonathon Snook ¹

To cite: Almilaji O, Smith C, Surgenor S, *et al*. Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia. *BMJ Open Gastro* 2020;**7**:e000403. doi:10.1136/bmjgast-2020-000403

Received 15 March 2020
Revised 30 March 2020
Accepted 8 April 2020

ABSTRACT

Objective To refine and validate a model for predicting the risk of gastrointestinal (GI) cancer in iron deficiency anaemia (IDA) and to develop an app to facilitate use in clinical practice.

Design Three elements: (1) analysis of a dataset of 2390 cases of IDA to validate the predictive value of age, sex, blood haemoglobin concentration (Hb), mean cell volume (MCV) and iron studies on the probability of underlying GI cancer; (2) a pilot study of the benefit of adding faecal immunochemical testing (FIT) into the model; and (3) development of an app based on the model.

Results Age, sex and Hb were all strong, independent predictors of the risk of GI cancer, with ORs (95% CI) of 1.05 per year (1.03 to 1.07, $p < 0.00001$), 2.86 for men (2.03 to 4.06, $p < 0.00001$) and 1.03 for each g/L reduction in Hb (1.01 to 1.04, $p < 0.0001$) respectively. An association with MCV was also revealed, with an OR of 1.03 for each fl reduction (1.01 to 1.05, $p < 0.02$). The model was confirmed to be robust by an internal validation exercise. In the pilot study of high-risk cases, FIT was also predictive of GI cancer (OR 6.6, 95% CI 1.6 to 51.8), but the sensitivity was low at 23.5% (95% CI 6.8% to 49.9%). An app based on the model was developed.

Conclusion This predictive model may help rationalise the use of investigational resources in IDA, by fast-tracking high-risk cases and, with appropriate safeguards, avoiding invasive investigation altogether in those at ultra-low predicted risk.

INTRODUCTION

Iron deficiency anaemia (IDA) is a common clinical problem, with an overall incidence in western populations approaching two cases per 1000 pa, and a considerably higher age-specific incidence in those over the age of 70 years.^{1,2} More than a quarter of men and postmenopausal women with IDA have significant underlying gastrointestinal (GI) pathology, and malignancy is by far the most important cause, found in 8%–10% of cases.^{3–5} IDA is an important indicator of GI cancer, particularly cancer of the right colon, as it often

Summary box

What is already known about this subject?

- ▶ Gastrointestinal (GI) cancer is the cause of iron deficiency anaemia (IDA) in 8%–10% of adult men and postmenopausal women.
- ▶ The risk of GI cancer in IDA is influenced by age, sex and Hb.
- ▶ Faecal immunochemical testing (FIT) in IDA may be of value in identifying underlying GI cancer.

What are the new findings?

- ▶ Age, sex and Hb are confirmed as strong predictors of the risk of GI cancer in IDA.
- ▶ Mean cell volume is an additional independent predictor of the risk.
- ▶ In combination, these four predictors can identify 10% of the referred IDA population who are at ultra-low risk of GI cancer.
- ▶ FIT is predictive of GI cancer risk in high-risk individuals with IDA, though the sensitivity is low.
- ▶ An app can facilitate the use of the model in a clinical setting.

How might it impact on clinical practice in the foreseeable future?

- ▶ The predictive model may allow the use of investigational resources to be rationalised in IDA, by fast-tracking high-risk cases and, with appropriate safeguards, avoiding invasive investigation altogether in those at ultra-low predicted risk.
- ▶ The app is intended to facilitate the use of this model in a clinical setting.



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Department of Gastroenterology, Poole Hospital NHS Foundation Trust, Poole, UK

²Clinical Research Unit, Bournemouth University, Bournemouth, Dorset, UK

³Health Technology Assessment Group, University of Central Lancashire, Preston, UK

Correspondence to

Dr Jonathon Snook;
jonathon.snook@gmail.com

occurs before any other clinical pointer to the diagnosis.⁶

The IDA clinic at Poole Hospital is the point of referral for the many patients with IDA who have minimal or no symptoms to indicate the nature or location of the underlying cause of iron deficiency and for whom further assessment is felt to be warranted. Basic patient data have been collected since inception for the purpose of clinical care,

audit and service evaluation. The referral rate to the IDA clinic now exceeds 400 new patients per annum.^{2,7}

In view of the possibility of underlying GI cancer, it is current standard practice to advise urgent investigation of at-risk subjects with IDA, which in the first instance generally involves gastroscopy and colonoscopy/colonography to examine the upper and lower GI tract, respectively.⁸ These investigations are however expensive and labour intensive, and not entirely without risk of problems and complications, particularly in those with significant comorbidities. Furthermore, over 80% of investigations for IDA will not reveal significant pathology.

As individuals with IDA are likely to vary in their individual likelihood of malignancy, a simple but reliable preinvestigation predictor of GI cancer risk would help considerably with patient counselling. Risk stratification could also rationalise the use of resources, with prioritisation of high-risk subjects for fast-track investigation, and perhaps avoidance of invasive investigation altogether in particularly low-risk individuals.

Previous work by our group and others^{9,10} has demonstrated that three simple and objective clinical variables—age, sex and blood haemoglobin concentration (Hb)—appear to be independent predictors of underlying GI cancer in IDA. In the IDIOM (Iron Deficiency as an Indicator of Malignancy (IDIOM) study of an IDA cohort of 720, the combination of these variables was used to derive a score corresponding to the percentage probability of underlying GI malignancy, which ranged from less than 2% in low-risk subgroups to more than 20% in high-risk subgroups.¹⁰ These studies^{9,10} do however have the shortcomings that both were retrospective in design and lacked an a priori hypothesis, simply because there was insufficient evidence on which to base such a hypothesis.

The aims of the study reported here were threefold. First, to provide prospective validation of the independent variables identified in the original IDIOM study as predictors of underlying GI cancer, by analysing a much larger IDA cohort, and to determine whether mean cell volume (MCV) and iron studies (transferrin saturation/serum ferritin) might prove to be additional predictors of risk. Second, to undertake a pilot study to explore whether faecal immunochemical testing (FIT) for small quantities of human haemoglobin in faecal specimens can improve risk stratification still further. The rationale for this hypothesis is that chronic low-grade blood loss from the tumour bed is assumed to be the major factor contributing to the development of IDA in subjects with GI cancer. Third, to develop an app for use in the clinical setting to provide an instant assessment of GI cancer risk following the input of simple clinical data.

METHOD

Validation study

The first part of the study involved a detailed assessment of clinical data for subjects referred for assessment

in the Poole IDA clinic with confirmed iron deficiency by standard laboratory criteria (transferrin saturation <15% and/or serum ferritin concentration less than the lower limit of the reference interval for the laboratory at the time) who were assessed between 2004 and 2018 inclusive,² incorporating some cases included in a previous report.¹⁰ Cases presenting in 2004–2016 formed the training dataset, while those presenting in 2017–2018 provided the validation dataset. Developing the model using the training dataset was carried out in 2018, before receiving the validation dataset.

The final datasets included age at presentation and sex, blood test results (Hb, MCV and iron studies) and the diagnostic findings on standard investigation of the upper and lower GI tract. Data sets were complete for age, sex, Hb, MCV and presence/absence of GI malignancy. As results were available for *both* transferrin saturation and serum ferritin in only 36.8% of the study population, iron deficiency was analysed as a dichotomous variable, being ‘severe’ (arbitrarily defined as a transferrin saturation <10% and/or a serum ferritin <10 µg/L) or ‘non-severe’ (criteria for severe deficiency not met).

Anonymised data were analysed to assess whether the five clinical parameters could usefully predict the likelihood of GI malignancy on subsequent investigation. Data preparation involved cleaning the data by checking and correcting any unusual values, removing duplicate entries and retaining only the first record for any patient referred more than once to the IDA clinic. A training dataset was used to derive the prediction model, which was then tested on a validation dataset. As this was a secondary analysis of anonymised data, formal Research Ethics approval was not required for this element of the study.

Logistic regression models were run for each of the predictors separately, with GI cancer as the outcome. When any significant association was found between a predictor and GI malignancy ($p < 0.05$), this predictor was added to a multivariable logistic regression model. Smoothed scatter plot, Cook’s distance and standardised residual errors, variance inflation factor, Akaike information criterion, analysis of variance χ^2 test, pseudo R^2 and the Hosmer-Lemeshow test were used to check the validity of the fitted logistic regression model and the goodness of fit.¹¹

To assess the performance of the fitted model derived from the training dataset, we examined how well it predicted GI cancer in the validation dataset. Cut-off metrics^{12,13} were used to assess performance, because traditional evaluations such as overall accuracy were not appropriate¹⁴ in view of the small percentage of participants with GI malignancy in the study. A classification cut-off probability (decision threshold) was identified using the training data, in which a value above that cut-off indicates the presence and a value below the absence of GI cancer. The prediction model was then tested on the validation dataset using this cut-off. Three optimal prediction cut-offs were selected :

1. Cut-off 1: the highest cut-off at which the negative predictive value (NPV) remains 100%. NPV is the number of negative cases that were correctly classified divided by the total number of negative cases predicted.¹⁵ This cut-off identifies subjects who are at ultra-low risk of GI cancer.
2. Cut-off 2: at which geometric mean (G mean) of sensitivity and specificity is highest.¹⁶ G mean is calculated from the formula: $\sqrt{(\text{sensitivity} * \text{specificity})}$.^{17 18}
3. Cut-off 3: the lowest cut-off at which the positive predictive value (PPV) remains in the upper quartile (ie, the point below which 75% of PPVs lie). PPV is the number of positive cases that were correctly classified divided by the total number of positive cases predicted.¹⁵ This cut-off identifies patients who are at high risk of GI cancer.

Receiver operating characteristic (ROC) was used to compare and visualise the effectiveness of the predictive model at separating positive and negative classes according to each cut-off.¹⁹

FIT pilot study

In brief, 80 subjects were prospectively identified who fulfilled all of the following criteria: (1) confirmed IDA, (2) high GI cancer risk based on age and Hb (70 years or over and <100 g/L, respectively)¹⁰ and (3) listed for investigation with gastroscopy and colonoscopy/colonography. Each was invited to provide a faecal sample for FIT prior to invasive investigation, using the HemaScreen SPECIFIC kit (Alpha Laboratories, Eastleigh, UK); the manufacturer's published analytical detection limit for this test is 50 µg Hb/g faeces.²⁰ FIT analysis was undertaken without knowledge of the outcome of GI investigation.

App development

To simplify utilisation of the prediction model in clinical settings, a web-based application was developed. R (V.3.6.1), RStudio (V.1.2.5001), R Shiny and DT packages were used to run the statistical analysis and to build the app.

RESULTS

Validation study

Over 2800 subjects with iron deficiency were seen in the IDA clinic during the study period. Excluding those in whom investigations were not completed due to patient preference, frailty or concurrent illness, and those whose records were incomplete, left 2390 subjects for detailed analysis. For the validation study, there were 1879 in the training dataset and 511 in the validation dataset.

The total study group comprised 1528 women and 862 men (a sex ratio of 1.8), with a median age of 71 years (IQR: 59–79 years) and mean (SD) values for Hb and MCV of 103 (17.4) g/L and 80.0 (9.1) fL, respectively. The arbitrary criteria for severe iron deficiency were met by 57% of the study population. GI carcinoma was

identified in 200 individuals in the study group, giving an overall prevalence of 8.4%. Of those, 172 (86%) were in the lower GI tract, and of those, 140 (81%) were in the right colon.

Comparison of the training and validation datasets revealed marginally higher values for mean Hb (102 g/L vs 106 g/L, $p < 0.001$) and mean MCV (79.4 fL vs 82.2 fL, $p < 0.001$) in the latter. This is consistent with changes in the characteristics of our IDA population over time reported elsewhere.² There were otherwise no significant differences between the training and validation datasets for any of the key variables.

Analysis of the training dataset confirmed that age, sex and Hb were all strong, independent predictors of the risk of GI cancer. MCV was also predictive though there was greater variability, resulting in a wider CI. There was no significant relationship with the results of iron studies. The final multiple binary logistic regression model was therefore constructed according to the formula: $\ln(\text{GI_cancer}) \sim \beta_0 + \beta_1 \text{age} + \beta_2 \text{sex} + \beta_3 \text{Hb} + \beta_4 \text{MCV}$. Statistical assessment of validity and goodness of fit of the logistic regression model based on the criteria outlined in the Method section was satisfactory.

The ORs (95% CI, p value) for the four predictive variables were as follows:

- ▶ Age: 1.05 per year (1.03 to 1.07, $p < 0.00001$).
- ▶ Sex: 2.86 for men (2.03 to 4.06, $p < 0.00001$).
- ▶ Hb: 1.03 for each g/L reduction (1.01 to 1.04, $p < 0.0001$).
- ▶ MCV: 1.03 for each fL reduction (1.01 to 1.05, $p < 0.02$).

The ROC curve for the training dataset shows the true positive rate on Y axis (sensitivity) and false positive rate on X axis (1-specificity), along with the three optimal cut-offs described in the Method section (figure 1). Using the

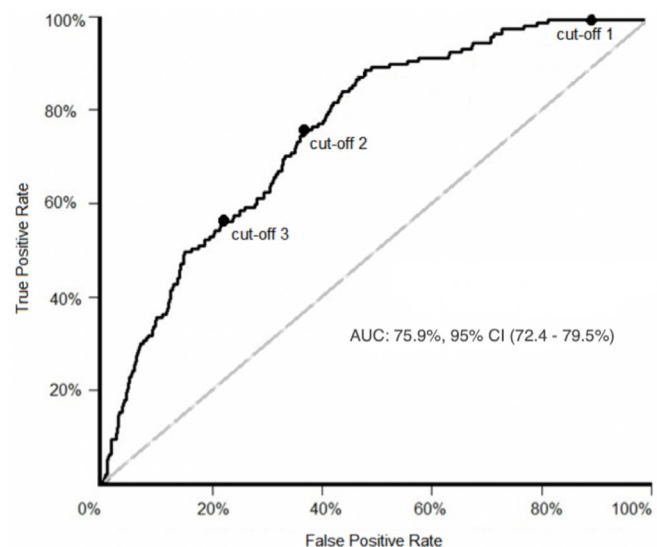


Figure 1 Receiver operating characteristic curve for the training dataset, showing the three optimal cut-off points defined in the text: cut-off 1=1.5%, cut-off 2=7.4%, cut-off 3=11.1%. AUC, area under curve.

Table 1 Characteristics of the three optimal cut-off points for predicted probability of GI cancer, as applied to the training and validation datasets

| Threshold | Criterion | | Training dataset (%) | Validation dataset (%) |
|------------------------------|-------------------------|----------------------|-----------------------|------------------------|
| Optimal cut-off 1 (1.5%) | NPV=100% | Sensitivity (95% CI) | 100.0 (97.7 to 100.0) | 100.0 (91.8 to 100.0) |
| | | NPV (95% CI) | 100.0 (98.0 to 100.0) | 100.0 (95.1 to 100.0) |
| Optimal cut-off 2 (7.4%) | G mean | Sensitivity (95% CI) | 75.8 (68.3 to 82.3) | 79.1 (63.9 to 89.9) |
| | | Specificity (95% CI) | 63.2 (60.9 to 65.5) | 67.7 (63.3 to 71.9) |
| Optimal cut-off 3 (11.1%) | PPV in highest quartile | Sensitivity (95% CI) | 56.1 (47.9 to 63.9) | 60.5 (44.4 to 75.0) |
| | | PPV (95% CI) | 18.6 (15.2 to 22.4) | 23.9 (16.2 to 32.9) |

GI, gastrointestinal; G mean, geometric mean; NPV, negative predictive value; PPV, positive predictive value.

regression model to calculate predicted GI malignancy risk, cut-off 1 (risk 1.5%) was able to stratify about 10% of both cohorts into an ultra-low risk subgroup. Cut-off 2 (risk 7.4%) maximised G mean in the training dataset (69.2%; 95% CI 21.8% to 219.9%) and gave a comparable value in the validation dataset (73.2%; 95% CI 27.4% to 195.6%), with closely overlapping CIs and similar values for sensitivity and specificity. Cut-off 3 (risk 11.1%) stratified about 25% of both cohorts into a high risk subgroup. These results (summarised in [table 1](#)) demonstrate that the model is robust in predicting the risk of underlying GI cancer in a new IDA dataset collected in a different time period.

The striking effect of combining the predictive variables on predicted risk is displayed in heat-map format

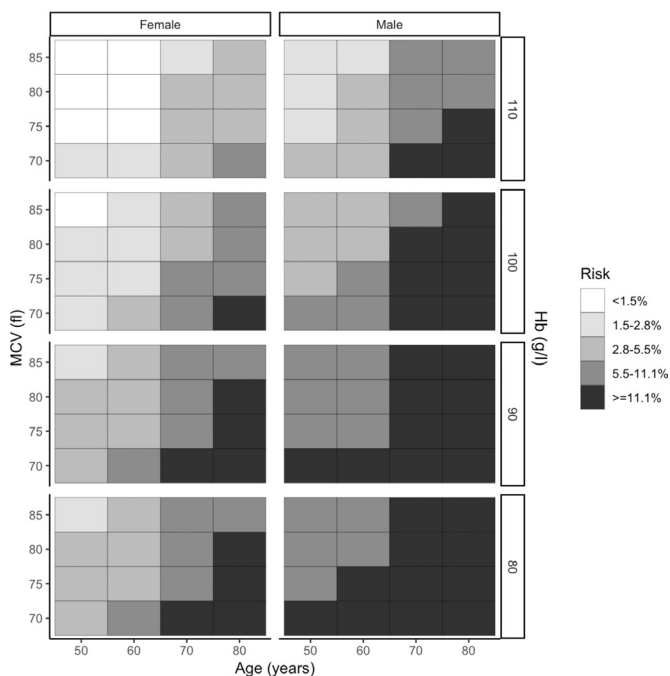


Figure 2 Heatmap showing the probability of gastrointestinal (GI) cancer in the overall IDA cohort (n=2390) according to age, sex, blood haemoglobin concentration (Hb: g/L) and mean cell volume (MCV: fL). The darker the box, the higher the GI cancer risk—as shown on the risk key. The risk ranges are based on positive predictive value quartiles, with the lowest quartile divided in two. IDA, iron deficiency anaemia.

in [figure 2](#). This demonstrates the high risk in all older men with IDA regardless of haematology findings, and the extremely low risk in younger women with marginal anaemia and a normal MCV. None of the individuals with a risk predicted by the model of less than 1.5% proved to have GI cancer on investigation—accounting for 10% of the whole cohort.

FIT pilot study

A total of 62 subjects at predicted high risk of GI malignancy returned an adequate faecal sample for FIT analysis and completed their scheduled investigations. Of these 17 (27.4%) proved on subsequent investigation to have a GI cancer (upper GI - 2, right colon - 14, left colon - 1). A summary of the results is shown in [table 2](#) - FIT positivity was associated with GI malignancy (OR=6.6, 95% CI 1.6 to 51.8), and this significant association persisted after adjustment for the IDIOM score variables of age, sex, Hb and MCV. However, the sensitivity of FIT for GI cancer was low at 23.5% (95% CI 6.8% to 49.9%), and this only increased to 26.7% (95% CI 7.8% to 55.1%) with exclusion of the upper GI cancers.

App development

An app (Predict GI Cancer in IDA) was developed based on the model. This generates an estimate of GI cancer risk (with 95% CI) following the insertion of data for the four key variables: age, sex, Hb and MCV. The whole process takes just a few seconds, which lends itself to use in busy clinical settings, and our intention is to make the app freely available following Medicines and Healthcare Products Regulatory Agency approval and CE marking. A screenshot from the app is shown in [figure 3](#).

Table 2 Distribution of gastrointestinal (GI) cancers by faecal immunochemical testing (FIT) result in 62 subjects with IDA at predicted high risk

| | GI cancer | |
|-----------------|-----------------|-----------------|
| | Negative (n=45) | Positive (n=17) |
| FIT result | | |
| Negative (n=56) | 43 | 13 |
| Positive (n=6) | 2 | 4 |

IDA, iron deficiency anaemia.

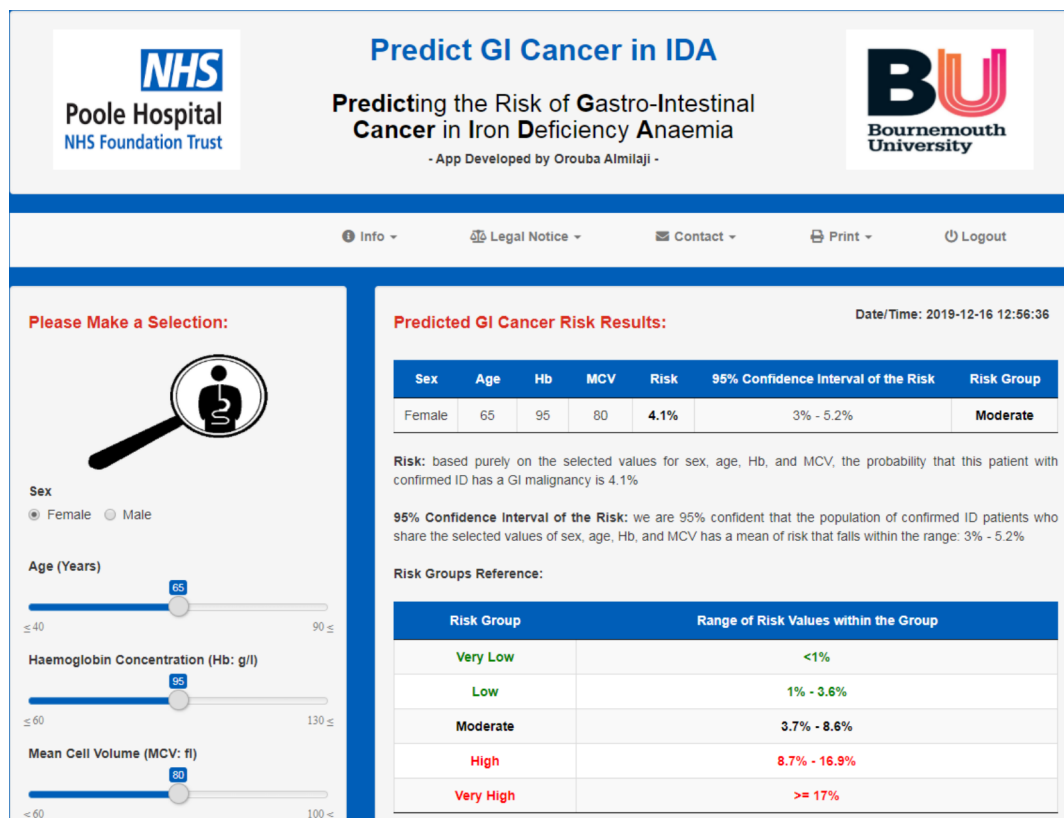


Figure 3 A screenshot from the app Predict GI Cancer in IDA.

DISCUSSION

IDA is a problem commonly encountered in clinical practice, and the prevalence of underlying GI cancer in IDA is the primary justification for urgent investigation.^{3–8} Bidirectional endoscopy (BDE), combining gastroscopy and colonoscopy in the same session, is generally accepted as the most efficient method of assessing the GI tract unless there are clear clinical clues as to the cause.⁷ It does however carry a small but significant risk of complications, particularly in the elderly and those with major comorbidities, and it is important to consider the risk–benefit ratio for the investigation of IDA on an individual case basis.

BDE is also labour intensive, taking up to an hour to complete for each patient, yet over 90% of procedures for IDA will not reveal malignancy. Because it is common, IDA is a major drain on investigational resources, accounting for a substantial proportion of the workload in many endoscopy units, with estimates in the region of 20% of all diagnostic examinations.² Any manoeuvre to safely reduce the number of necessary investigations has the potential to make a substantial positive impact on both costs and waiting times.

There is therefore the need for a simple and reliable pretest predictor of the risk of underlying malignancy that is sufficiently discriminating to be clinically useful for patient-centred counselling. Effective risk stratification is a potentially useful clinical tool for two reasons. First, it allows the identification of a high-risk subgroup who warrant accelerated investigation and can be advised

accordingly. Second, it reveals individuals at very low risk who are unlikely to benefit from invasive investigation and may wish to make a considered decision not to proceed. The development of an app means that GI cancer risk can be computed in a few seconds, with obvious benefit in busy clinical settings.

The findings of this study have limitations. First, the predicted GI cancer risk is in all cases greater than 0% and less than 50%. Second, while GI cancer is the most important cause of IDA, it is not the only one, and we know from previous work that the model is not useful in predicting the likelihood of these other causes.¹⁰ For these two reasons, the model can never be more than a guide to the need for invasive investigation. Finally, while large the study is based on a single-centre experience, raising the question of universal applicability. Work is underway to address this by validating the model on a totally independent external IDA dataset.

The study reported here builds on previous reports from our group and others^{9 10} by confirming in a much larger IDA cohort that age, sex and Hb are all strong independent predictors of the risk of GI cancer. It also reveals an independent relationship with MCV – this has not been evident in previous analyses^{9 10} apart from a single report on a very small cohort²¹, and has perhaps emerged in this study because of the substantially larger cohort size.

The predictive value of age and sex is not unexpected, given that the incidence of the major GI malignancies rises steeply after the age of 70 years, particularly in



men.^{22 23} It may be that Hb is predictive of GI cancer risk simply because the nature of the pathology means that GI malignancy is disproportionately more likely than the other (non-malignant) causes of IDA to lead to greater degrees of anaemia.

The explanation for the effect of MCV on risk is less clear. It might perhaps reflect either chronicity or severity of the depletion of body iron stores in those with underlying GI cancer. Although the analysis of iron studies does not support the latter explanation, ferritin and transferrin saturation are surrogate markers of iron stores and may be influenced by other factors. Serum ferritin in particular is an acute phase protein and may therefore be spuriously high in individuals with malignancy.

IDA is a particular challenge in the elderly,²⁴ as this is the age group with the highest prevalence of IDA, and the highest risk of underlying GI cancer.² However, it is also the age group at highest risk of complications from invasive investigation or from subsequent surgery if required—and debatably the least to gain from intervention. Management planning in this situation needs to be made on a case-by-case basis, and while only one element of the risk–benefit equation, an accurate prediction of GI cancer risk can only help the individual concerned to reach the right decision.

One of the striking findings of the study is the identification of subgroups with a very low GI cancer risk. Indeed, in the 10% of the total cohort with a predicted risk of less than 1.5%, no GI cancers were found. It is important to note that this includes some postmenopausal women, as shown in [figure 2](#). The finding is unlikely to be the result of referral bias, as younger women with mild anaemia are the IDA subgroup least likely to be referred unless there was some other reason for suspecting GI disease, for example, a strong family history of GI cancer.

It is important to stress that ‘low-risk’ does not equate to ‘no risk’ and that additional fail-safes need to be incorporated before advocating a no investigation policy for low-risk subgroups, a process known as diagnostic safety netting.²⁵ The first safety net for ‘low risk’ IDA is ensuring a full and sustained haematological response to a course of iron replacement therapy. This should already be standard practice and has been shown to predict a very low risk of missed pathology following BDE in those with IDA.²⁶

A second potential safety net is testing for tiny quantities of blood in a faecal sample using FIT. The development of FIT is undoubtedly a major step forward in the risk assessment of patients in primary care presenting with lower GI symptoms and in screening programmes for colorectal cancer (CRC) such as the NHS England Bowel Cancer Screening Programme.^{27–30} It has a greater sensitivity for CRC (the most common GI cancer underlying IDA) than guaiac-based testing for faecal occult blood^{31 32} and has been shown to be of some predictive value for GI cancer in the IDA population without clinical risk scoring.^{27 33 34} The situation might be analogous to established practice in the diagnosis of pulmonary

embolism, where it is accepted that those with a low clinical probability score *and* a low test result (for d-dimer) have such a vanishingly low risk that further investigation is not warranted.³⁵

The pilot study reported here demonstrates that in a high-risk IDA subgroup FIT can predict the presence of CRC, but the sensitivity of 26.7% is disappointingly low. Numbers are obviously small, but this suggests that FIT may not be a particularly helpful adjunct to the IDIOM score in predicting GI cancer risk, at least at the 50 µg Hb/g faeces detection threshold. It may be that FIT at a lower detection threshold might improve the sensitivity for CRC in IDA without an unacceptable fall in specificity, although a recent meta-analysis demonstrates only a marginal improvement in sensitivity on reducing the FIT threshold from ≥ 30 to 10 µg Hb/g faeces, despite more than doubling the number of positive results.³⁶

The low sensitivity found here may at first sight seem surprising, but it is important to bear in mind that while right-sided lesions account for about 35% of all CRCs, the figure is over 80% for the subgroup presenting with IDA.² Concerns have been raised about the sensitivity of FIT for right-sided CRC,³³ and two recent real-world studies have confirmed that this is an issue, reporting that about 10% of all CRCs had a FIT of less than 10 µg Hb/g faeces, most of these being right-sided tumours presenting with IDA.^{28 29} An analysis of quantitative FIT results revealed median concentrations of 41.6 and 286.8 µg Hb/g faeces for right-sided (n=17) and left-sided (n=23) CRCs, respectively (p<0.03).²⁹

A recent systematic review of CRC detection by FIT in IDA cohorts yielded five studies with a sensitivity of 0.82 (95% CI 0.68 to 0.90), though most were small, and the evidence quality was poor with a high risk of bias.²⁷ Further research in this area is warranted, but the provisional conclusion must be that a negative FIT does not reliably exclude CRC in the context of IDA. Following on from this, it may be safest to regard IDA and FIT as complementary indicators of the possibility of underlying CRC.

In conclusion, this study has extended previous observations, confirming that the simple and objective criteria of age, sex and Hb are strong and independent predictors of the risk of underlying GI cancer in subjects with IDA, and the additional benefit of incorporating MCV into the risk stratification model. It has demonstrated that in combination these variables can identify 10% of the study population who are at ultra-low risk. The development of an app based on this model adds practical value in a clinical setting.

Contributors OA, AC, PT, EW and JS conceived and designed this study, and CS and SS collected the data. OA analysed the data and drafted the initial manuscript, and JS is the guarantor. All authors made significant contributions to the subsequent revision of the paper and approved the final version prior to submission.

Funding (1) Faecal immunochemical testing kits funded by Poole Hospital Gastroenterology Research Fund. (2) PhD studentship (OA) jointly funded by Poole Hospital Gastroenterology Research Fund and Bournemouth University. (3) AC

part-funded by the National Institute for Health Research Applied Health Research Collaboration. The views expressed are those of the authors and not necessarily those of the NIHR or Department of Health and Social Care.

Disclaimer The views expressed are those of the authors and not necessarily those of the NIHR or Department of Health and Social Care.

Competing interests SS and EW have received honoraria for speaking at educational meetings sponsored by Pharmacosmos.

Patient consent for publication Not required.

Ethics approval A pilot study to explore the potential role of FIT—IDIOM-3 (ISRCTN No 18342140)—was undertaken with Research Ethics approval (IRAS No 201759).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request. Data available from corresponding author.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID ID

Jonathon Snook <http://orcid.org/0000-0002-3172-2722>

REFERENCES

- 1 Camaschella C. Iron-Deficiency anemia. *N Engl J Med* 2015;372:1832–43.
- 2 Stone H, Almilaji O, John C. The dedicated iron deficiency anaemia clinic – a fifteen-year experience.
- 3 Rockey DC. Gastrointestinal tract evaluation in patients with iron deficiency anemia. *Semin Gastrointest Dis* 1999;10:53–64.
- 4 Liu K, Kaffes AJ. Iron deficiency anaemia: a review of diagnosis, investigation and management. *Eur J Gastroenterol Hepatol* 2012;24:109–16.
- 5 Zhu A, Kaneshiro M, Kaunitz JD. Evaluation and treatment of iron deficiency anemia: a gastroenterological perspective. *Dig Dis Sci* 2010;55:548–59.
- 6 Rai S, Hemingway D. Iron deficiency anaemia—useful diagnostic tool for right sided colon cancers? *Colorectal Dis* 2005;7:588–90.
- 7 Surgenor SL, Kirkham S, Parry SD, et al. The development of a nurse-led iron deficiency anaemia service in a district general Hospital. *Frontline Gastroenterol* 2014;5:219–23.
- 8 Goddard AF, James MW, McIntyre AS, et al. Guidelines for the management of iron deficiency anaemia. *Gut* 2011;60:1309–16.
- 9 James MW, Chen C-M, Goddard WP, et al. Risk factors for gastrointestinal malignancy in patients with iron-deficiency anaemia. *Eur J Gastroenterol Hepatol* 2005;17:1197–203.
- 10 Silva AC, Sheppard ZA, Surgenor SL, et al. Clinical risk factors for underlying gastrointestinal malignancy in iron deficiency anaemia: the IDIOM study. *Frontline Gastroenterol* 2014;5:237–42.
- 11 Agresti A. *Categorical data analysis*. New Jersey: John Wiley & Sons Inc, 2002.
- 12 Ferri C, Hernández-Orallo J, Modroui R. An experimental comparison of performance measures for classification. *Pattern Recognit Lett* 2009;30:27–38.
- 13 Chen JJ, Tsai C-A, Moon H, et al. Decision threshold adjustment in class prediction. *SAR QSAR Environ Res* 2006;17:337–52.
- 14 Afonja T. *Accuracy paradox*. Towards Data Science, 2017.
- 15 Tharwat A. Classification assessment methods. *Applied Computing and Informatics* 2018;12.
- 16 Kubat M, Holte R, Matwin S. Learning when negative examples abound. In: van Someren M, Widmer G, eds. *ECML 1997*. Heidelberg: Springer LNCS, 1997: 1224. 146–53.
- 17 Yu H, Hong S, Yang X, et al. Recognition of multiple imbalanced cancer types based on DNA microarray data using ensemble classifiers. *Biomed Res Int* 2013;2013:1–13.
- 18 Lin W-J, Chen JJ. Class-imbalanced classifiers for high-dimensional data. *Brief Bioinform* 2013;14:13–26.
- 19 Fernández A, García S, Galar M, et al. *Learning from imbalanced data sets*. 1st Edn. Switzerland:Springer, 2018.
- 20 Alpha Laboratories UK. hema-screen specific fit, 2018. Available: <https://www.alphalabs.co.uk/media/productfile/file/h/e/hemascreen-specific-2012.pdf> [Accessed 15 Dec 2019].
- 21 Capurso G, Baccini F, Osborn J, et al. Can patient characteristics predict the outcome of endoscopic evaluation of iron deficiency anemia: a multiple logistic regression analysis. *Gastrointest Endosc* 2004;59:766–71.
- 22 Hamilton W, Lancashire R, Sharp D, et al. The importance of anaemia in diagnosing colorectal cancer: a case-control study using electronic primary care records. *Br J Cancer* 2008;98:323–7.
- 23 Cancer Research UK. Cancer Incidence Statistics, 2008 - 2010. Available: <http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/uk-cancer-incidence-statistics> [Accessed 15 Dec 2019].
- 24 Burton JK, Yates LC, Whyte L, et al. New horizons in iron deficiency anaemia in older adults. *Age Ageing* 2020;49:309–18.
- 25 Almond S, Mant D, Thompson M. Diagnostic safety-netting. *Br J Gen Pract* 2009;59:872–4.
- 26 Sahay R, Scott BB. Iron deficiency anaemia—how far to investigate? *Gut* 1993;34:1427–8.
- 27 Lee MW, Pourmorady JS, Laine L. Use of fecal occult blood testing as a diagnostic tool for clinical indications: a systematic review and meta-analysis. *Am J Gastroenterol* 2020;115:662–670.
- 28 Mowat C, Digby J, Strachan JA, et al. Impact of introducing a faecal immunochemical test (fit) for haemoglobin into primary care on the outcome of patients with new bowel symptoms: a prospective cohort study. *BMJ Open Gastroenterol* 2019;6:e000293.
- 29 Chapman C, Bunce J, Oliver S, et al. Service evaluation of faecal immunochemical testing and anaemia for risk stratification in the 2-week-wait pathway for colorectal cancer. *BJS Open* 2019;3:395–402.
- 30 Koo S, Neilson LJ, Von Wagner C, et al. The NHS bowel cancer screening program: current perspectives on strategies for improvement. *Risk Manag Healthc Policy* 2017;10:177–87.
- 31 Smith A, Young GP, Cole SR, et al. Comparison of a brush-sampling faecal immunochemical test for hemoglobin with a sensitive guaiac-based fecal occult blood test in detection of colorectal neoplasia. *Cancer* 2006;107:2152–9.
- 32 Selby K, Levine EH, Doan C, et al. Effect of sex, age, and positivity threshold on fecal immunochemical test accuracy: a systematic review and meta-analysis. *Gastroenterology* 2019;157:1494–505.
- 33 Nakama H, Zhang B, Fattah AS, et al. Colorectal cancer in iron deficiency anemia with a positive result on immunochemical fecal occult blood. *Int J Colorectal Dis* 2000;15:271–4.
- 34 Cilona A, Zullo A, Hassan C, et al. Is faecal-immunochemical test useful in patients with iron deficiency anaemia and without overt bleeding? *Dig Liver Dis* 2011;43:1022–4.
- 35 Kearon C, de Wit K, Parpia S, et al. Diagnosis of pulmonary embolism with D-dimer adjusted to clinical probability. *N Engl J Med* 2019;381:2125–34.
- 36 Kim NH, Yang H-J, Park S-K, et al. Does low threshold value use improve proximal neoplasia detection by fecal immunochemical test? *Dig Dis Sci* 2016;61:2685–93.