Bournemouth University

The Importance of Identifying Iron Deficiency Anaemia in the Early Detection of Colorectal Cancer

Orouba Almilaji

This Thesis is submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy Degree (PhD)

Medical Science and Public Health Department, Faculty of Health and Social Sciences, Bournemouth University, United Kingdom

Gastroenterology Unit, Poole Hospital, University Hospitals Dorset NHS Foundation Trust, United Kingdom

January 2022

Abstract

The importance of identifying iron deficiency anaemia in the early detection of colorectal cancer

Orouba Almilaji

Colorectal cancer (CRC) is common and carries a relatively poor prognosis. The strong relationship between tumour stage at diagnosis and survival is the basis of the English Bowel Cancer Screening Programme (BCSP) and highlights the importance of early diagnosis. Iron deficiency anaemia (IDA) is also common. About 10% of cases in males and post-menopausal females are due to underlying gastro-intestinal (GI) cancer, most commonly CRC - and IDA is often the first manifestation. This thesis examines the detailed relationship between IDA and CRC.

Chapters 4 and 5 describe the analysis of four large IDA datasets, confirming the prevalence of GI cancer, and demonstrating that cancer risk can be predicted from four simple, objective clinical indicators. This IDIOM model proved robust on internal and external validation. This research is valuable for patient counselling, targeting the investigation of high-risk individuals and (perhaps) avoiding invasive investigation in ultra-low risk cases.

Chapter 6 outlines the analysis of a subset with recurrent IDA, suggesting that the subsequent risk of GI cancer is higher in those who were incompletely investigated the first time around.

Chapter 7 describes the development of the *IDIOM App*. This is a freely available web-tool which allows cancer risk in IDA to be calculated within seconds, lending itself to clinical usage.

Chapters 8 and 9 report the analysis of a large CRC database, demonstrating that diagnosis through the IDA pathway (1) generates as many cases as the BCSP; (2) identifies a distinct sub-population with a predominance of right-sided lesions; and (3) like the BCSP, is associated with a favourable tumour stage profile.

The findings suggest that identifying iron deficiency anaemia could play an important role in the early diagnosis of CRC.

List of contents

List of figures

List of supplementary figures

List of tables

List of supplementary tables

Abbreviations

AJCC	American Joint Committee on Cancer
ASG	American College of Gastroenterology
AUC	The Area Under the Curve
BCSP	Bowel Cancer Screening Program
BDE	Bidirectional Endoscopy
BSG	British Society of Gastroenterology
BU	Bournemouth University
BUCRU	Bournemouth University Clinical Research Unit
CA	Cancer Antigen
CEA	Carcinoembryonic Antigen
CRC	Colorectal Cancer
CRUK	Cancer Research UK
CT	Computed Tomography Scan
EGD	Esophagogastroduodenoscopy
FAP	Familial Adenomatous Polyposis
FBC	Full Blood Count
Fe^{2+}	Ferrous Iron
Fe^{3+}	Ferric Iron
FIT	Faecal Immunochemical Test
FPN	Ferroportin
gFOBt	Guaiac Faecal Occult Blood Test
GI	Gastrointestinal
Hb	Haemoglobin Concentration
IDA	Iron Deficiency Anaemia
IDIOM	Iron Deficiency as An Indicator Of Malignancy
MAP	Mutyh Associated Polyposis
MCH	Mean Corpuscular Haemoglobin
MCV	Mean Red Cell Volume
MDD	Medical Device Directive
MDR	Medical Device Regulation

MHRA	Medicines and Healthcare Products Regulatory Agency
MRI	Magnetic Resonance Imaging
NCRAS	National Cancer Registration and Analysis Service
NCRI	National Cancer Research Institute
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NPV	Negative Predictive Value
ONS	Office for National Statistics
OR	Odds Ratio
PET	Positron Emission Tomography
PHE	Public Health England
PPV	Positive Predictive Value
RBC	Red Blood Cells
RDW	Red Cell Distribution Width
ROC	Receiver Operating Characteristic
UICC	Union for International Cancer Control
UKCA	UK Conformity Assessed Marking
WHO	World Health Organization

Acknowledgement

First and foremost, I would like to express my sincere appreciation and gratitude to my mentor and supervisor; Dr. Jonathon Snook. During every step in this PhD path, he was always present to academically encourage, guide, read, evaluate (and critically re-evaluate), and make sure that the outputs of this research are up to the standards of the doctoral degree. Without his continuous bright advice, profound scientific background, extended clinical expertise, and devoted dedication to this research project, I truly could not achieve what I was able to accomplish during this four-years PhD.

I also would like to acknowledge:

1. Professor Peter Thomas, my first ex-supervisor, who has retired unexpectedly during the first Covid-19 lockdown. Peter has scrutinized every little detail in the statistical analysis during the first two years, overseen carefully all the needed ethical approvals for this research, and taught me how to trust my formal statistical training. And for all that I will be always thankful for him.

2. Dr. Sharon Docherty who has voluntarily stepped in as supervisor for me when Peter has left. Without having Sharon as a new supervisor, I would not recover from the academic setback because of Covid-19. She kindly cared for my well-being as much as my academic progress and considerably never save an effort to make certain that this integrated format thesis meets the requirements of the doctoral degree before its submission. I really hope that I had her on board as supervisor since the beginning of my PhD.

3. Dr. Elisabeth Williams, Dr. Sally Parry, Ms Carla smith, Ms Lisa Cranham and every member of the Gastroenterology Department at Poole hospital for giving me the opportunity to work with them as an honorary research

fellow during my PhD, and for their simulating discussions, and close assessments and reviews of this PhD project.

4. Mrs Lesley Hutchins, the research commercialisation manager at Bournemouth University, who has made the registration of the *IDIOM App* a reality. Indeed, no words can express my gratitude to Lesley for her genuine, and highly respected commitment to not wasting three years of a tremendous work on the app development and writing the technical documentation during which I felt hopeless when trying to register the app.

5. Dr. Vegard Engen, Ms Rosemary Collins, Mr Derek Kelly, BU Innovations Limited (BUI), Dr Juan Campos-Perez, The Research Development and Support (RDS) at Bournemouth University, The IT service at Bournemouth University, and Oxentia Oxford's Global Innovation Consultancy, for their valuable suggestions, and support when writing the technical documentation of the app and/or certifying it.

6. Dr. Katie Sheehan, my line manager at King's College London, who has hired me three time to work as part-time research assistant on interesting projects during my PhD. This experience has helped significantly to improve my programming skills further by working on big data, apply new statistical methods, and interact and learn from amazing new researchers from outside the network of my PhD project.

7. My friends, Wafa, Halima, Eman, Suad, Rahaf, Rania, Natalie, Nour, Amal, Gina, and all my colleagues in office 313 at Royal London House.

8. The *fans in denial* whom main contribution to the scientific research will be remembered one day through their bullying, fabricating stories, spreading gossip, and trying to set other researchers up. A whole and complete set of integrity, intelligence, passion for research, education, and hard work is what should define a real academic researcher. At best, few in this fans'

"*gang"* would be qualified as careerists. Their constant intimidation was an additional reason that immensely enthused me to finish this PhD at the utmost level I could reach; and to become more vigilant, empathetic, and supportive towards those who are subject to covert racism, alienation, and discrimination in academia.

Finally, I would like to dedicate this thesis to:

- My mother: who always managed patiently to put her agreeable pretending face when I used to explain to her about all the statistical methods I am applying, and for encouraging me to finish this PhD.

- My husband: who was and still; a real motivation for any woman to learn, have a degree, and work.

- My daughter: my dearest friend, and a software engineer to be! Regardless, what I had to face with during this PhD journey, I kept striving to finish because I was shy to fail while she was watching.

- Aleppo, Syria: the home I was born, bred in, and honoured to belong to.

- All cancer patients/victims and their families -including mine- who inspired me to undertake this research.

Author's Declaration

This thesis compromises only the author own work, derived from original research undertaken after the date the author initially registered with the University. Six original research articles are submitted as integrated elements of this thesis - chapters 4 to 9 inclusive.

In all these papers the PhD candidate was the first and lead author, contributing to at least of 75% of the substantive content of each paper. For each jointly-authored paper, a breakdown of each individual's contribution can be found underneath the contribution subtitle.

As well as these 6 articles, various publications resulted from the PhD research. A summary of the publications per chapter is presented next:

- **Chapter 1, 2 and 3**: Introduction, background, and methods

  The general PhD research topic about the association between gastrointestinal cancer and iron deficiency anaemia, summary of the research data, and initial results were accepted as poster presentations at Bournemouth University:

  1. Almilaji, O., Thomas, P., & Snook, J. The importance of identifying Iron deficiency anaemia in the early detection of colorectal cancer. *10th Annual Postgraduate BU conference*. Mar 2018, Bournemouth University (Appendix Ia1).
  2. Almilaji, O., Thomas, P., & Snook, J. The importance of identifying Iron deficiency anaemia in the early detection of colorectal cancer (update). *BU PGR Live Exhibition*. Dec 2018, Bournemouth University (Appendix Ia2).

  And orally presented by the thesis author at:

  1. Gastrointestinal Departmental Clinical Governance Meeting on 25th of Oct 2019, Poole Hospital.

2. Gastrointestinal Departmental Clinical Governance Meeting on 18th of Dec 2018, Poole Hospital.
3. *CoPMRE Visiting Faculty Day* on 9th of May 2018, Executive Business Centre, Bournemouth University.

- **Chapter 4**: Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia.

  The whole chapter is submitted as part of the integrated thesis format and was published:

  Almilaji, O., Smith, C., Surgenor, S., Clegg, A., Williams, E., Thomas, P., & Snook, J. 2020. Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia. *BMJ Open Gastroenterology* 2020;7**:**e000403. doi: 10.1136/bmjgast-2020-000403.

  Material from the same chapter has been accepted as poster presentation (Appendix Ia3) at the British Society of Gastroenterology (BSG) Annual Meeting, Glasgow Jun 2019, and published:

  Almilaji, O., Thomas, P., & Snook, J. 2019. PWE-042 Predicting the risk of gastrointestinal cancer in iron deficiency anaemia. *Gut* 68**:**A192.

- **Chapter 5**: Broad external validation of a multivariable risk prediction model for gastrointestinal malignancy in iron deficiency anaemia.

  The whole chapter is submitted as part of the integrated thesis format and was published:

  Almilaji, O., Webb, G., Maynard, A., Chapman, T. P., Shine, B., Ellis, A. J., Hebden, J., Docherty, S., Williams, E. J., & Snook, J. 2021. Broad external validation of a multivariable risk prediction model for gastrointestinal malignancy in iron

deficiency anaemia. *Diagnostic and prognostic research*, *5*(1), 23. https://doi.org/10.1186/s41512-021-00112-8.

Material from the same chapter has been accepted as poster presentations (Appendix Ia4 and Ia5) and published at the conference websites:

1. Almilaji, O., Webb, G., Chapman, T. P., Williams, E. J., Shine, B. S. F., Ellis, A. J., Docherty, S., & Snook, J. Internal and External validation of the IDIOM score for predicting the risk of gastrointestinal malignancy in iron deficiency anaemia. *National Cancer Research Institute (NCRI) Virtual Showcase.* 2 -3 Nov 2020.
2. Almilaji, O., Webb, G., Chapman, T. P., Williams, E. J., Shine, B. S. F Ellis, A. J., Docherty, S., & Snook, J. External validation of the IDIOM score for predicting the risk of gastrointestinal malignancy in iron deficiency anaemia. Cancer Research UK *(CRUK) Early Detection of Cancer Conference.* 6 -8 Oct 2020, Online.

- **Chapter 6**: Modelling the episodes of care for IDA patients in a secondary-care centre using continuous-time multistate Markov chain.

  The whole chapter is submitted as part of the integrated thesis format and was published:

  Almilaji, O. 2022. Modelling the Episodes of Care for IDA Patients in A Secondary-Care Centre Using Continuous-Time Multistate Markov Chain. Saudi J Gastroenterol. 28:115-21. DOI: 10.4103/sjg.sjg_387_21

  Material from the same chapter was accepted as poster presentations (Appendix Ia6 and Ia7) at:

  1. Almilaji, O. 2021. P2985 - Modelling the episodes of care for IDA patients in a secondary-care centre using continuous-time multistate Markov chain. *American College of*

Gastroenterology (ACG) Annual Scientific Meeting. 22-27 Oct 2021.

2. Almilaji, O., Docherty, S., & Snook, J. Modelling the episodes of care for IDA patients in a secondary-care centre using continuous-time multistate Markov chain. (ID: 3556). *Virtual NCRI Festival: Making cancer research better together*. 8-12 Nov 2021.

ACG abstract is published:

Almilaji, O., S1313 Modelling the Episodes of Care for IDA Patients in a Secondary Care Centre Using Continuous-Time Multistate Markov Chain, The American Journal of Gastroenterology: October 2021 - Volume 116 - Issue - p S605 doi: 10.14309/01.ajg.0000778784.81979.c6.

- **Chapter 7**: The development of a web-based application to predict the risk of gastrointestinal cancer in iron deficiency anaemia; the IDIOM app.

The whole chapter is submitted as part of the integrated thesis format and was published:

Almilaji, O., Engen, V., Snook, J., & Docherty, S. 2022. The development of a web-based application to predict the risk of gastrointestinal cancer in iron deficiency anaemia; the IDIOM app. Digital 2022, 2, 104-119. https://doi.org/10.3390/digital2010007.

Material from the same chapter has been presented orally by at:

1. Virtual workshop between Bournemouth University (UK) and Northeastern University (China), IDIOM: A Software Medical Device to Predict the risk of GI cancer in IDA, on 4th Dec 2020, online.
2. FHSS Research Seminar Series: Rehabilitation and Sports Science Department organised for the MSc Nutrition and

Behaviour students on 5th of Feb 2020, Bournemouth University.

3. FHSS Research Seminar Series: Bournemouth University Clinical Research Unit (BUCRU) Digital Health on 20th of Feb 2019, Royal London House, Bournemouth University.

Material from the same chapter has been accepted as oral and poster presentation (Appendix Ia8) at BSG Campus (online), Jan 2021, and published:

Almilaji, O., Engen, V., Snook, J., & Thomas, P., 2021. The development of a web-based application to predict the risk of GI cancer in IDA. https://youtu.be/DRjenumKdhY. *Gut* 70:A37-A38.

- **Chapter 8**: Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia

The whole chapter is submitted as part of the integrated thesis format and was published:

Almilaji, O., Parry, S., Docherty, S., & Snook, J. 2021. Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia. *Scientific Reports*. 13055. https://doi.org/10.1038/s41598-021-92623-z

Material from the same chapter has been presented as poster at BSG Campus (online), Jan 2021 (appendix Ia9) and published:

Almilaji, O., Parry, S., Thomas, P. and Snook, J., 2021. Downstaging of right-sided colorectal cancer diagnosed through iron deficiency anaemia. https://youtu.be/R7_llJjRBEo. *Gut*;70:A190.

- **Chapter 9**: Colorectal cancer and the blood loss paradox

The whole chapter is submitted as part of the integrated thesis format and was published:

> Almilaji, O., Parry, S., Docherty, S., & Snook, J. 2021. Colorectal cancer and the blood loss paradox. Frontline Gastroenterology. October 2021. doi: 10.1136/flgastro-2021-101959.

Additional Outputs

There are additional Four PhD outputs are not included as integral parts of this thesis but derived from the PhD research. These are:

- A confidential report of 400 pages that contains proprietary information and intellectual property to register the *IDIOM App* with the Medicines and Healthcare products Regulatory Agency (MHRA) as first-class software medical device. The report was written and is updated by the thesis author and include the administrative information of the app, the technical documentation, reference to the relevant standards, the code, and the declaration of conformity. The technical documentation includes the general description of the app and its intended use, the determination whether it is a medical device, the classification discussion, the relevant conformity assessment, and the development process documents of the *IDIOM App*. The development process documents of the *IDIOM App* include purpose, scope, definitions, development planning, requirements analysis, implementation, deployment, clinical evaluation, risk analysis, release and label, and the usability assessment questionnaires. The code includes the analysis code, the app interface, the server code, and the deployment relevant setup code. The report can be accessed and viewed at BUCRU. A copy of the declaration of the conformity is included in this thesis (Appendix Ib).

- The certified app itself which is developed, deployed, managed, and maintained by the thesis author. A complete citation of the app is given next.

  Almilaji, O., Thomas, P., & Snook, J. 2020. The IDIOM App (ver 1.0): A Web-based Application to Predict the Risk of Gastrointestinal Cancer in Iron Deficiency Anaemia. BU Innovations Limited (BUI), UK. [Software Medical Device, Class I]. Available from: https://www.predict-gi-risk-in-ida.com. [Accessed on 14/12/2021].

- Published Paper:

  Stone, H., Almilaji, O., John, C., et al. 2020. The dedicated iron deficiency anaemia clinic: A 15-year experience. Frontline Gastroenterology. doi: 10.1136/flgastro-2020-101470.

  Though the thesis author was a co-author in this paper, responsible for the statistical analysis based on the PhD datasets. However, because the PhD candidate was not the first author, the paper is not submitted as a part of this integrated thesis, nor should it be assessed as such.

- Paper in preparation:

  Almilaji O, Docherty S, Snook, J. 2021. Lessons learned from using a secondary data: Dealing with duplicates that share the same characteristics.

In summary, the up-to-date outputs of this PhD project are:

- 1 certified clinical support decision software medical device app.
- 1 technical documentation relates to the app development project.
- 8 original research articles (7 published, and 1 in preparation).
- 9 peer-reviewed published conference papers.
- 7 oral presentations.

Among these outputs, only 6 published original research articles were submitted as part of this integrated thesis.

Chapter 1 : Introduction

1.1 Outline

This chapter presents the context, objectives and importance of the study. Then it outlines the structure of this thesis. Finally, it summarises the impact of Covid-19 on delaying the timeline of PhD project.

1.2 Study scope

Gastrointestinal (GI) cancer is the most common cause of cancer death worldwide in 2020 (WHO 2021). Colorectal cancer (CRC) is most common type of GI cancer; the 2nd most common cause of cancer death in the UK, accounting for 10% of all cancer deaths in 2018; and the only GI cancer type that has its national screening programme. When diagnosed at its earliest stage, more than 91% people with CRC will survive their disease for five years or more, compared with 10% people when the disease is diagnosed at the latest stage (ONS 2019). Also, when diagnosed at its earliest stage, 65% people with stomach cancer - another type of GI cancer - will survive their disease for five year or more, compared with around 24% when the disease is diagnosed at the latest stage (ONS 2019).

With the aim of reducing the former mortality rate by both earlier detection of CRC and removing polyps which, if left untreated may advance to cancer (Almilaji et al. 2021a), the Bowel Cancer Screening Programme (BCSP) has been introduced in England (PHE 2019). The BCSP included two arms, though the second is currently suspended:

- Biennial home-based Guaiac faecal occult blood test (gFOBt) which was introduced in 2006 and changed in 2019 to a more sensitive faecal immunochemical tests (FIT) in the English BCSP. The test is offered to patients aged 60–74 years and if it is found positive, patients are offered a colonoscopy.

- Bowel scope test, which was introduced in 2013 to patients aged 55 years and involves a one-off invitation for a flexible sigmoidoscopy. If pre-cancerous polyps are found, the patient is offered a colonoscopy.

Regular bowel cancer screening has been shown to reduce the CRC mortality by about 15% (Koo et al. 2017), probably because 63% of CRC cases detected through the BCSP are diagnosed at early stages (stage I or II) (NCRAS 2016). However, because only about 10% of the total CRCs in the UK are detected through this BCSP route (NCRAS 2016; Koo et al. 2017; Nelson et al. 2017), the efficiency of this programme is still low. Predominantly, BCSP is inefficient in detecting the right-side CRCs cases, as 74% of the screen-detected CRCs in the UK are found in the left side, and only 26% are found in the right (Braun et al. 2016).

A potential approach to the earlier diagnosis of GI cancer in general, and CRC in particular is through the detection of iron deficiency anaemia (IDA). Much research has highlighted the association between GI cancer in both the upper and lower GI tract and IDA (Rocky and Cello 1993; Silva et al. 2014; Hung et al. 2002; Shahriari-Ahmadi et al. 2017; Goddard et al. 2011; Wijayasekara et al. 2016). Patients with IDA are at increased risk for in particular right-sided CRC (Goodman and Irvin 1993; Stebbing and Nash 1995; Alexiusdottir et al. 2012; Edna et al. 2012; Schop et al. 2019; Tokunaga et al. 2019; Niv et al. 2005).

Since IDA can be caused by chronic bleeding GI lesions (Niv et al. 2005), many of which are malignant, an essential component of patient evaluation is by bidirectional endoscopy (BDE). Especially in the evaluation of unexplained iron deficiency, BDE is indicated to rule out any neoplasia (Hunt and Faigel 2002). The endoscopic examination of the GI tract in IDA patients could provide an opportunity for the diagnosis of early-stage GI cancer at a curable stage.

However, these examinations are negative in a high proportion of cases without overt bleeding (Cilona 2011). And, in practice, the performance of a high-quality examination is varied, with noticeably an unacceptable high rate of failure to diagnose cancer (Beg et al. 2017).

Often worse quality of cleansing of the right colon (Brenner et al. 2010), also, because of their flat morphology; right side CRCs are much more difficult to be picked up by these examinations than left side CRCs (Heresbach et al. 2008). Unfortunately, as a result, right-sided CRCs are commonly detected in more advanced stages than left side CRCs (Baran et al. 2018).

More importantly, BDE is labour-intensive, time-consuming procedure, and might cause complications particularly in the elderly (Almilaji 2020).

gFOBt and FIT are simple and non-invasive investigations that have already been shown to be of some value in identifying IDA patients due to underlying GI malignancy (Nakama 2001; Chowdhury 2014; Kim et al. 2017; Selby et al. 2019). FIT in particular has proven to have a high negative predictive value (NPV) in IDA (Ayling et al. 2019), in which NPV refers here to the probability that IDA patients with a negative FIT result truly do not have the GI cancer. Hence, FIT could be useful to prioritize BDEs (Cilona 2011).

1.3 Study rationale

Due to the prevalence of IDA in the population, IDA has a considerable impact on referrals for urgent investigation for suspected cancer (Snook et al. 2021). Previous research showed that a substantial number of referrals are inappropriate according to guidelines, and may lead to considerable workload, financial implications, and be detrimental to patient health (Shaw et al. 2008; Mankodi et al. 2010).

Therefore, it is important to identify sub-groups of IDA patients who are either at high risk of GI cancer, and so warrant fast-track BDE; or at such low risk that they could reasonably be managed without the need for invasive investigation. And, to assess whether FIT might allow further stratification of IDA patients prior to invasive investigation.

As GI cancer is frequently found in patients with IDA, the burden of diagnosing and treating anaemia has shifted gradually from primary care physicians and haematological specialists to gastroenterologists (Zhu et al. 2010). This shift has led to the opening of the one of the first dedicated IDA clinics in the UK by the Gastroenterology Unit in Poole hospital (Wijayasekara et al. 2016).

To assist with the patient counselling and prioritization of investigational resources, and by analysing data (n=720) generated by this particular clinic, the Iron Deficiency as an Indicator Of Malignancy (IDIOM) study (Silva et al. 2014), has been conducted. Using multivariable analysis of the predictive value of age, sex, mean red cell volume (MCV), haemoglobin concentration (Hb), and iron studies for the risk of underlying GI malignancy, the study concluded that only age, sex, and Hb were associated with the risk of GI cancer.

The study demonstrated that three simple and objective clinical variables; sex, age, Hb can be combined to provide a useful GI cancer risk stratification model (IDIOM score) for IDA patients. By combining these three risk factors, 12% of the study population were identified at particularly low risk of GI cancer (low risk defined as risk <2%), and 16% as high risk (risk >20%). The model was validated in 2016, using new data (n= 643) from the same clinic and no other predictive clinical variables were identified (Wijayasekara et al. 2016).

Along with the IDIOM score study, there were only other three published studies (Capurso et al. 2004; Ho et al. 2005; James et al. 2005) that developed a multivariable risk prediction model to predict the risk of GI cancer in IDA. The sample size for these studies was 98, 148, 695 respectively.

Though age was a universal positive predictor of the GI cancer risk, as expected, in all previous models as well as the IDIOM score, the results were conflicting with regard to the other predictors; sex, Hb, iron studies, and MCV. One explanation for these inconsistent results might be caused by the small size of the studies especially the first two (Capurso et al. 2004; Ho et al. 2005). Another explanation could be the forcing of the quarters or dichotomous classification of continuous predictor variables in the predictive model. Age and Hb for instance, were coded into categories in both the IDIOM score model, and James et al. (2005) study.

Categorization of continuous data should be avoided in the statistical analysis as it leads to information loss, underestimation of the extent of variation in outcome between groups, and concealment for any non-linearity in the relation between the variable and outcome (Altman and Royston 2006). Further work was obviously required to refine the IDIOM score.

By utilising larger different clinical datasets collected from the IDA clinic in Poole Hospital and from other secondary centres in the UK, this PhD research comprises a series of statistical investigations aiming to provide a detailed appraisal of the relationship between GI cancer and IDA, an assessment of the potential benefit of screening for IDA on the prognosis of CRC, and to ultimately build a digital decision support tool that enables simple access to health professionals in a busy clinical setting.

1.4 Study objectives

The aims of the PhD project are:

1. Examining whether it is possible to add additional clinical and laboratory predictor(s) that could improve the performance of the IDIOM model.

2. Examining the generalisability of the IDIOM model by validating it using temporal dataset from the same IDA clinic.

3. Examining whether FIT can improve risk stratification still further.

4. Examining the transportability of the IDIOM prediction model by validating it using independent external data from other centres in the UK.

5. Stratifying the IDA patients into GI risk groups according to the externally/internally validated IDIOM model.

6. Investigating the proper methods to estimate the delay time spent between the two consecutive referrals, where an IDA patient did not complete investigations at the first referral, but at the second is diagnosed with positive GI cancer.

7. Developing an automated decision-support tool in which anonymised, individual, patient data is entered, and GI cancer risk is calculated based on the IDIOM model.

8. Examining whether CRC diagnosed through the detection of IDA are associated with an earlier tumour stage.

9. Examining the association between prior blood count test events and CRC stage.

10. Examining whether BCSP and IDA pathways identify different CRC sub-populations (left-sided and right-sided CRC).

1.5 Thesis outline

Since this PhD project compromises several individual proposed studies, each addressing different research question(s) and analysing different dataset(s), the format of the integrated thesis was deemed appropriate.
The thesis is composed of ten chapters. Chapters 4 to 9 are all already published as original research articles and appear in the corresponding journal format.

Also, due to the permittable integrated thesis format, each of these chapters have self-contained components including their own abstract, keywords, abbreviations, introduction, literature review, background, aim(s), methods, data, results, discussion, figures, tables, and references. Because of the nature of publishing research articles, there were elements of the research study that not published as integral parts of these published original research articles. However, these elements were published separately in the journals as supplementary information and will be added at the end their chapters.

The incorporation of publication-style chapters has led to some components that repeated / overlapped with parts of the other sections in the current thesis format. Upon a wishful acceptance of this thesis, only an online link to the article that was published as non-open access will be added under the chapter title (chapter 9).

The PhD thesis starts with this chapter; then sequentially moves into the following chapters:

Chapter 2. Background

This chapter explains the relevant concepts and background to this research project, some of these concepts were skipped totally in the papers. These notions relate to anaemia, iron deficiency anaemia, iron metabolism, IDA

diagnosis, and treatment. Also, the chapter defines many terms related to the GI cancer including diagnosis, staging, and treatment.

Chapter 3. General methods

Being a PhD project in applied medical statistics, the description of the statistics analysis and software development in all the published papers was very extensive. Nonetheless, this chapter outlines the used datasets and main statistical approach per each objective. The summary of the software development project. The data management and ethical approvals for all the individual studies in this PhD project.

Chapter 4. Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia.

After introducing the relevant concepts to the PhD project, and the summary of the methods, the analysis started in this chapter by addressing three objectives from this PhD. These are examining whether the existing IDIOM score model can be improved by adding a new predictor based on using larger dataset from the same IDA clinic at Poole hospital (objective 1). Then checking the model accuracy and performance, by validating it internally using data from the same IDA clinic but from a different time period (objective 2). Finally, based on the new validated model, the chapter examined whether the IDA risk stratification can be improved by adding the result of FIT to the model (objective 3).

Chapter 5. Broad external validation of a multivariable risk prediction model for GI malignancy in iron deficiency anaemia

To apply the model which was developed in chapter 4 with confidence to different populations, it must be tested, and amended in case of poor performance, using external validation dataset (objective 4). For this purpose, two datasets from Oxford and Sheffield were used. Only after this, it was possible then to finalise the stratification of the IDA patients into GI risk groups

(objective 5) according to the successful externally and internally validated IDIOM model.

Chapter 6. Modelling the episodes of care for IDA patients in a secondary-care centre using continuous-time multistate Markov chain

Using the risk groups which were proposed in the last chapter, this chapter examined whether being stratified in ultra-low risk or very-high risk group at the earlier episode of care is associated with being diagnosed with positive GI cancer at the following episode of care. This chapter investigated the proper methods to estimate the delay time spent between the two consecutive referrals, where an IDA patient did not complete investigations at the first referral, but at the second is diagnosed with positive GI cancer (objective 6).

Chapter 7. The development of a web-based application to predict the risk of gastrointestinal cancer in iron deficiency anaemia; the IDIOM app

Having built a risk prediction model for GI cancer in IDA patient, internally and externally validated it successfully in the last two chapters. It was then possible to build a tool that makes use of this model and enables simple access to healthcare professionals (objective 7). And indeed, using R the *IDIOM App* was developed and certified. Since there are, currently, limited resources on developing standalone software medical devices this chapter aimed to document the development of a web-based clinical support decision app in an academic setting.

Chapter 8. Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia

The results from earlier chapters confirmed that IDA is good marker for GI cancer. But given that the efficiency of BCSP is still low in detecting the right-side CRCs, and patients with IDA are at increased risk for right-sided CRC,

this chapter examined whether being diagnosed through the IDA pathway can lead to downstage the right-sided CRC and accordingly improve its prognosis (objective 8). And examined the association between prior blood count test event and CRC stage (objective 9).

Chapter 9. Colorectal cancer and the blood loss paradox

The results from earlier chapter confirmed that IDA pathway can lead to downstage the right-sided CRC, having a prior blood count test is related to detecting CRC at early-stage, and recommended a blood count test in the at-high risk population. To further support this recommendation, this chapter compared the clinical characteristics of patients with CRC diagnosed via BCSP and IDA pathways, and examined whether these pathways indeed identify different CRC sub-populations (objective 10).

Chapter 10. General Discussion

Since only online links will be retained for each published chapter in the final thesis body, this chapter recapped the key findings of each chapter extensively, explained its different limitations, critically and cautiously interpreted the results in relation to the objectives, discussed the new knowledge this project has generated, and any potential future research plans.

1.4 Impact of Covid-19 pandemic on research

Without warning, Covid-19 pandemic has struck in 2020 causing loss, grief and disruption to many lives. For the first time during this PhD project, the boundary between personal and academic life was diminished. Access to the office PC and desk at the university was no longer an option. Being a sole carer for an elderly mother with advanced heart disease, and a mother and academic support for a young woman who is at the last year at college

studying and preparing to go to the university, and all being locked in a very small apartment for months has introduced additional challenges to maintaining the momentum of this research, and resulted in increased stress of balancing home and research demands simultaneously.

There were time periods in the pandemic in which research was significantly stalled or even stopped totally. These periods including four admissions for my mother to the hospital because of her deteriorating health during the pandemic. A waiting period to arrange a desk, a chair, and two large monitors in order to make a progress with external validation analysis, and many other situations. Because of the pandemic, about extra one year was added to the PhD initial 3 years period timeline.

Despite all these barriers which imposed by the pandemic, there were some positive prospects of growing academically associated with pandemic. These include the participation in different conferences which pivoted to a virtual setting allowing reduced cost of registration, and more inclusive participation. The publication in academic journals which waived article processing charges (APCs), and/or granted extended time to the submission and doing the required corrections requested by the reviewers. Moreover, resorting to the scientific community via social media during the pandemic lockdown was an opportunity to nurture and develop professional relationships and collaborations with different researchers across the world.

Most importantly, this pandemic has demonstrated strongly how important is the scientific research, including this PhD project. As a contrary to the title of this section, this research has actually a positive impact on the pandemic which has put unprecedented additional pressure on the Trust. Exceptionally, one of this PhD project's outputs, namely; The *IDIOM App* was introduced into general clinical practice by the IDA clinic at Poole Hospital to triage patients during the Covid-19 lockdown, when investigational resources were extremely limited.

Chapter 2 : Background

2.1 Outline

This chapter provides relevant background to the main concepts and terms used in this research, dividing it into two main aspects - IDA and GI cancer. The first aspect includes the general concept of anaemia, IDA, iron metabolism, IDA diagnosis, and treatment. And the second aspect involves defining GI cancer and related terms - diagnosis, staging and treatment.

2.2 Iron deficiency anaemia (IDA)

2.2.1 Overview

The Hb molecule is the combination of a protein (globulin) and non-protein portion (haem) in red blood cells (RBCs), and is responsible for transporting oxygen from the lungs to the body's tissues and returning carbon dioxide from the tissues back to the lungs (Patel et al. 2021). Anaemia is a serious worldwide public health problem that affects both developing and developed countries (Snook et al. 2021). The word “anaemia” came from the ancient Greek “anaimia” which means ‘lack of blood’ (Johnson-Wimbley and Graham 2011). The WHO (2011) definition of anaemia is based on the haemoglobin (Hb) concentration, in which anaemia is described as a condition in which the number of red blood cells and / or the Hb within them is lower than normal. The definition of anaemia is Hb level below: 110 g/dL for children under 5 years, 115 g/L for children 6–11 years, 120 g/L for children 12–14 years, 130 g/L for adult males, 120 g/L for adult non-pregnant females and 110 g/L for adult pregnant females (WHO 2011).

Anaemia can be developed as a result of malabsorption, malnutrition, chronic / or acute bleeding, systemic inflammation, decline in endogenous erythropoietin (EPO) production, or a reduction in bone marrow response to

endogenous erythropoietin (Busti et al. 2018). However, the most common causes are haemoglobinopathies, nutritional deficiencies and infectious diseases such as HIV, malaria and tuberculosis (WHO 2011). The most recent published statistics by WHO shows that 50% of all anaemia cases worldwide are attributable to iron deficiency (WHO 2011). IDA is a form of anaemia due to the lack of sufficient iron to form normal red blood cells. IDA is typically a microcytic anaemia, defined by a decreased MCV (De Franceschi et al. 2017), and it affects 2.36 billion people worldwide (Disease and Injury Incidence and Prevalence Collaborators 2016). Common symptoms of IDA are headache, fatigue, pale skin, and dyspnoea (Lopez et al. 2016).

2.2.2 Iron metabolism

Because iron studies are fundamental concepts to this PhD study, this section introduces the related process to them in the human body by summarising how iron is absorbed into the body, regulated, used, recycled, and stored. The section is collectively derived from published literature (Lopez et al. 2016; Polin et al. 2013; Camaschella 2017; NICE 2021; Conard et al. 1966; Raffin et al. 1974).

Absorption

Iron enters the body either as haem iron that is come from animal source or as non-haem iron that is from vegetables. Haem iron is more easily absorbed into the body because it is not affected by the many ligands in the food; moreover, it is directly taken up into enterocytes. Non-haem iron in the form "Fe^{3+}" (ferric iron) is reduced to the more easily absorbed form "Fe^{2+}" (ferrous iron) by stomach acids, ferric reductase, and duodenal cytochrome b (DCYTB). Also, in the intestinal lumen, pancreatic enzymes play a role in freeing haem from globulin. Absorbed iron (1–2 mg per day) enters the intracellular iron pool of enterocytes (intestinal epithelial cells) by a transporter called intestinal haem iron (HCP1). Most of the absorption happens in the duodenum and jejunum.

Recycling

Parallel to the absorption process is the process of recycling the iron inside the body through special type of white blood cells called macrophages. These cells digest and engulf microbes, cancer cells, foreign substances, cellular debris, and anything else that does not have proteins types specific to healthy body cells on its surface in a process called phagocytosis. In fact, the major source of iron in the body is provided by the macrophages that recycle iron from senescent red blood cells. Macrophages phagocytize and degrade damaged erythrocytes, uptake Hb, breakdown haem, predominantly to produce Hb in new erythrocytes.

Storage and usage

If the body does not require the iron, it is stored as protein ferritin mainly in the liver, or it may be transported to the bone marrow to make new red blood cells.

For the iron to be exported into the bloodstream, it must be transported by transmembrane protein that release the iron across the basolateral membrane of enterocytes and macrophages, called ferroportin (FPN). For the ferrous iron to be transported in the bloodstream it must act as an acceptor and hence back to the form Fe^{3+}. To facilitate this oxidisation chemical reaction, hephaestin and ceruloplamin assist.

Regulation

Iron release from macrophages, enterocyte, or iron storing hepatocytes is closely regulated by a peptide hormone produced by hepatocytes called hepcidin. Hepcidin binds to this sole cellular exporter ferroportin, and hence prevents iron entry into plasma. Hepcidin is stimulated by plasma iron and iron stores and stimulated by inflammation. Increased erythropoietic activity (the process of making red blood cells that is driven by decreased Oxygen levels in the circulation) inhibits hepcidin from binding to ferroportin through producing a protein hormone called erythroferrone. This hormone causes an increase in the amount of iron that is available for Hb synthesis. As a response

to hypoxia, the kidneys secrete a glycoprotein cytokine called erythopoietin to stimulate the bone marrow to make more red blood cells.

2.2.3 How IDA is diagnosed and treated

The cause of IDA can be broadly attributed to malabsorption of dietary iron and poor dietary intake, as well as several significant GI pathologies (Snook et al. 2021).

For people with suspected IDA, a full blood count (FBC) should be arranged (NICE 2015). If results of the full blood count show a low Hb without obvious explanation, ferritin levels should be checked (NICE 2015). In all people, a serum ferritin level of less than about 30 µg/L confirms the diagnosis of iron deficiency (NICE 2015) though the exact figure varies between laboratories. However, ferritin levels are challenging to interpret if infection or inflammation is present, as levels can be high even in the presence of iron deficiency, and its level may be less reliable in pregnancy (NICE 2015). Other blood tests such as transferrin saturation can be helpful if a false-normal ferritin is suspected (Snook et al. 2021).

Initial treatment of IDA should be with oral ferrous sulphate, fumarate, or gluconate, with intravenous iron, or (rarely) limited transfusion of packed red cells (Snook et al. 2021). Iron replacement therapy should not be deferred while awaiting investigations for IDA unless colonoscopy is imminent (Snook et al. 2021). Hb levels (full blood count) should be checked after to assess the patient's response to iron treatment (NICE 2021).

Bleeding, in general, is a known problem in cancer patients, related to local tumour invasion, abnormal tumour vasculature, or tumour regression (Johnstone and Rich 2018). Overt or occult bleeding is commonly prominent in GI cancers (Busti et al. 2018). Since IDA can be an outcome for a chronic low-grade blood loss from a malignant GI tumour (Almilaji et al. 2020), GI investigation on an urgent basis should be considered in adults with a new

diagnosis of unexplained IDA (Snook et al. 2021). In fact, per current guidelines in the UK, CRC referral should be arranged urgently using a suspected cancer pathway for an appointment within 2 weeks for people aged over 60 years with IDA (NICE 2015). And urgent referral should be considered for people with unexplained IDA aged under 50 years with rectal bleeding (NICE 2015). For people aged under 60 years with IDA and without rectal bleeding, quantitative faecal immunochemical tests should be offered (NICE 2015).

2.3 Gastrointestinal cancer (GI)

2.3.1 Overview

GI cancer is the most common cancer worldwide in 2020 (3.02 million new cases), affecting both men and women (WHO 2021). GI cancer refers to malignancies of the GI tract and its accessory organs of digestion such as the oesophagus, gallbladder, pancreas, stomach, and intestine (Tian and Hu 2020). The GI wall surrounding the lumen of the GI tract is made up of four layers of specialised tissues - from the lumen outwards: the mucosa (epithelium, lamina propria, and muscular mucosae); the submucosa; the muscularis propria (inner circular muscle layer, intermuscular space, and outer longitudinal muscle layer); and the serosa (Rao and Wang 2010).

The upper GI tract is constituted by the oral cavity and salivary glands, the oesophagus, stomach, and small intestine (duodenum, jejunum, and ileum) (Treuting et al. 2017). The lower GI tract comprises the appendix, caecum, colon, rectum, and anus (Treuting et al. 2017). The colon is approximately 1.5 metres in length, and composed of five parts - the caecum, ascending colon, transverse colon, descending colon and sigmoid (Gervaz et al. 2004).

This distinction between left-sided and right-sided colon is based on their embryological origins (Bufill 1990). The caecum, appendix, ascending colon,

hepatic flexure, and proximal two thirds of the transverse colon have originated from the midgut. Whereas distal one third of the transverse colon, splenic flexure, sigmoid colon, descending colon, and rectum originated from the hindgut (Bufill 1990; Imperial et al. 2018; Iacopetta 2002; Mik et al. 2017). Hence, right-sided CRCs arise from the epithelial tissue of the caecum, ascending colon, or proximal two thirds of the transverse colon and the left-sided CRCs arise from the descending, sigmoid, or distal one third of the transverse colon as illustrated in Figure 2-1.

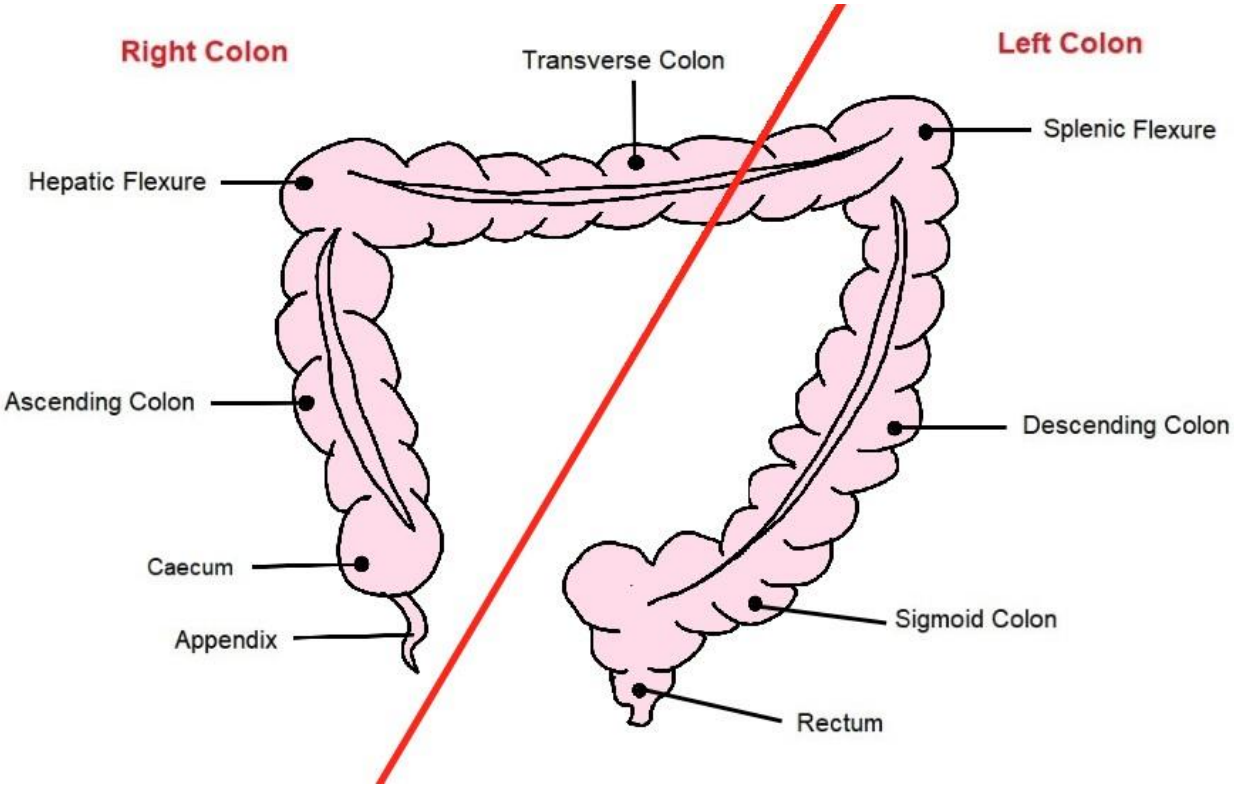


Figure 2-1 Right and left colon

Besides the difference in their origin, the tumours of right-sided CRCs and left-sided CRCs have different histology (Baran et al. 2018). Right-sided CRCs have flat morphology that is more difficult to detect with colonoscopy screening than left-sided tumours (Nitsche et al. 2016; Brenner et al. 2010; Gualco et al. 2006; Nawa et al. 2008). Beside their flat morphology, often worse quality of cleansing of the right colon (Brenner et al. 2010), and

because some “complete” colonoscopies do not evaluate the entire right colon (Baxter et al. 2009) right-side CRCs are commonly detected in more advanced stages than left side CRCs (Baran et al. 2018; Heresbach et al. 2008). In the UK, 74% of the screen-detected CRCs are found in the left side (Braun et al. 2016). And only 26% of these screen-detected CRCs are found in the right side (Braun et al. 2016).

2.3.2 GI cancer diagnosis

If GI cancer is suspected, different GI cancer investigations (Hiom 2015; PHE 2017; Rubin et al. 2018) may be performed. These include:

1. Genetic testing to check for family history. Examples of these genetic conditions include FAP (Familial Adenomatous Polyposis), Lynch syndrome, and MAP (MUTYH Associated Polyposis).
2. Endoscopy or esophagogastroduodenoscopy (EGD) to check upper GI tract.
3. Flexible sigmoidoscopy to check the lower part of the colon.
4. Colonoscopy to check the lower GI tract.
5. Lab tests to look for changes in the blood that could be signs of cancer, i.e., tumour marker such as cancer antigen (CA) 72-4, carcinoembryonic antigen (CEA) and cancer-related antigen 19-9 (CA 19-9).
6. Imaging studies (ultrasound, X-ray, magnetic resonance imaging (MRI), computed tomography scan (CT) scan, or positron emission tomography (PET) scan) to check for malignancies in the digestive system.
7. Biopsy to obtain a sample of abnormal tissue and analyse it for the presence of malignancy. This is a part of procedure 2, 3, 4, and 6.
8. Faecal immunochemical test (FIT).

2.3.3 GI cancer staging

There are different systems usually used to classify GI cancer tumours. These include the TNM system, the number staging system, and the grading system.

Tumour, nodes, and metastases (TNM) system

The American Joint Committee on Cancer (AJCC 1977) and the Union for International Cancer Control (UICC 2016) maintain the TNM classification system to stage many different types of cancer based on certain common standards. This overall stage is based on assigning the cancer in question a letter and number to describe the tumour (T), node (N), and metastasis (M) categories.

T describes the original (primary) tumour.

- TX: no information about the primary tumour, accordingly it cannot be evaluated.
- T0: no evidence of a primary tumour, accordingly it cannot be found.
- Tis: in situ cancer or pre-cancer. This means cancer cells are only growing in the layer of cells where they started and not into deeper layers.
- T1-T4: the tumour size and/or amount of spread.

N describes whether the cancer has spread to the nearby lymph nodes.

- NX: no information about the nearby lymph nodes, accordingly it cannot be evaluated.
- N0: nearby lymph nodes (regional) do not contain cancer.
- N1- N3: size, location, and/or the number of nearby affected lymph nodes.

M describes whether the cancer has spread (metastasized) to distant parts of the body

- M0: no distant cancer spread has been found.
- M1: cancer has been found to have spread to distant organs or tissues.

Number staging system

This is another commonly used staging system derived from Dukes classification (Dukes 1932). Stages are described with the Roman numerals I, II, III, and IV (National Health Service (NHS) 2018; Hawkes 2019). Stages, sometimes, are further subdivided using the letters A, B and C that imply certain characteristics found on work-up of the cancer.

- Stage 0: This is also known as "carcinoma in situ," or cancer has not spread.
- Stage I: The tumour is usually small and has not spread to nearby lymph nodes or major organs.
- Stage II: The tumour has grown, but has not spread into nearby tissue, and have spread into the lymph nodes.
- Stage III: The cancer is large and spread and more advanced and aggressive than Stage II.
- Stage IV: This also known as "secondary" or "metastatic" cancer.

Grading system

This system describes how the cancer cells compared with normal cells look under the microscope (National Health Service (NHS) 2018).

- Grade 1: well differentiated or low-grade is when the cancer cells look like normal cells and usually grow slowly.
- Grade 2: intermediate-grade or moderate is when the cancer cells look more abnormal and are slightly faster growing.
- Grade 3: poorly differentiated or high-grade is when the cancer cells look very different from normal cells and may grow rapidly.

2.3.4 Treatment

A combination of different treatments may be tailored to each patient based on different factors such as cancer stage, type and patient age. These treatments (National Health Service (NHS) 2019a; National Health Service (NHS) 2019b; American Cancer Society 2020) include:

- Chemotherapy: drugs to destroy cancer cells by stopping their ability to divide and grow.
- Targeted therapy: drugs that target the specific genes, proteins, or the tissue environment that contributes to cancer survival and growth.
- Surgery: traditional open procedure, or minimally invasive laparoscopic, or robotic procedure.
- Radiotherapy: high doses of radiation to shrink tumours and destroy cancer cells.

Chapter 3 : General methods

3.1 Outline

This chapter outlines the datasets used and the main statistical approach for each objective, a summary of the software development project, and the data management and ethical approvals for all the individual studies in this PhD project.

3.2 Study design

The PhD project consists of two major parts. These are

3.2.1 A series of statistical investigations

a. For anonymised clinical data collected retrospectively from existing patients records:

1. *IDA training dataset*: Patients (n=2295) referred to Poole Hospital IDA clinic between 2004 and 2016. The dataset includes age at presentation, sex, blood test results (Hb, MCV, and iron studies (Ferritin and /or Transferrin saturation)), and the diagnostic findings on standard investigation of the upper and lower GI tract.

2. *IDA validation dataset*: Patients (n=602) referred to Poole Hospital IDA clinic between 2017 and 2018. The dataset includes age at presentation, sex, blood test results (Hb, MCV and iron studies), and the diagnostic findings on standard investigation of the upper and lower GI tract.

3. *Oxford validation dataset*: Patients (n=1147) fast-track referred to two hospitals in Oxford between 2016 and 2019. The dataset includes age at presentation, sex, blood test results (Hb, MCV and iron studies), and

the diagnostic findings on standard investigations of the upper and lower GI tract.

4. *Sheffield validation dataset*: Patients (n=477) referred to dedicated IDA clinic in one hospital in Sheffield between 2013 and 2018. The dataset includes age at presentation, sex, blood test results (Hb, MCV and iron studies), and the diagnostic findings on standard investigation of the upper and lower GI tract.

5. *CRC dataset*: Patients (n=1258) diagnosed with CRC in the Gastroenterology Department in Poole hospital between 2010-2016 and presented through different pathways. The dataset includes age at diagnosis; sex; Hb at presentation; presentation pathway (IDA, BCSP or symptomatic); tumour stage; tumour number; histology; and location(s).

6. *CRC_IDA detailed dataset*: Patients (n=171) diagnosed with CRC in the Gastroenterology Department in Poole hospital and presented through IDA pathway between 2010- 2016. The dataset includes age at diagnosis; sex; the date and Hb result at presentation; tumour stage; tumour number, histology; location(s); whether a blood count had been checked in the 3 years prior to presentation; and if so, the date and Hb result for the last blood count.

7. *Referrals Dataset*: Patients (n=83) referred to Poole Hospital IDA clinic between 2004 and 2018, each with more than one visit to the IDA clinic (a total of 168 episodes of care). The data includes sex, Hb, MCV, iron studies, the diagnostic findings on standard investigation of the upper and lower GI tract, date of the visit(s) to the IDA clinic, age at each visit, and Indicator of the GI investigations' completion.

b. For anonymised clinical data collected prospectively following a simple universal intervention (FIT assessment):

FIT dataset: Patients (n=80) who have been invited to Poole Hospital IDA clinic between 2017 and 2019 to provide a faecal sample for FIT prior to invasive investigation. The dataset includes age at presentation, sex, blood test results (Hb, MCV and iron studies), the diagnostic findings on standard investigation of the upper and lower GI tract (and side), and the FIT results.

The setting for all these individual cohort studies was secondary care. The total number of the centres in this PhD project is 3; Dorset (Poole Hospital), Oxford, and Sheffield. Patients in all these studies were adults referred to secondary care and who went on to be investigated for IDA. Confirmation of IDA depended on the local lab in each centre. All subjects underwent standard GI investigation for IDA; comprising exclusion of coeliac disease, OGD and an adequate colonic examination – either by CT colonography or colonoscopy.

For the purpose of transparency, a visual demonstration of the overlapping between these datasets is illustrated next in Figure 3-1, in which n stands for the number of patients:

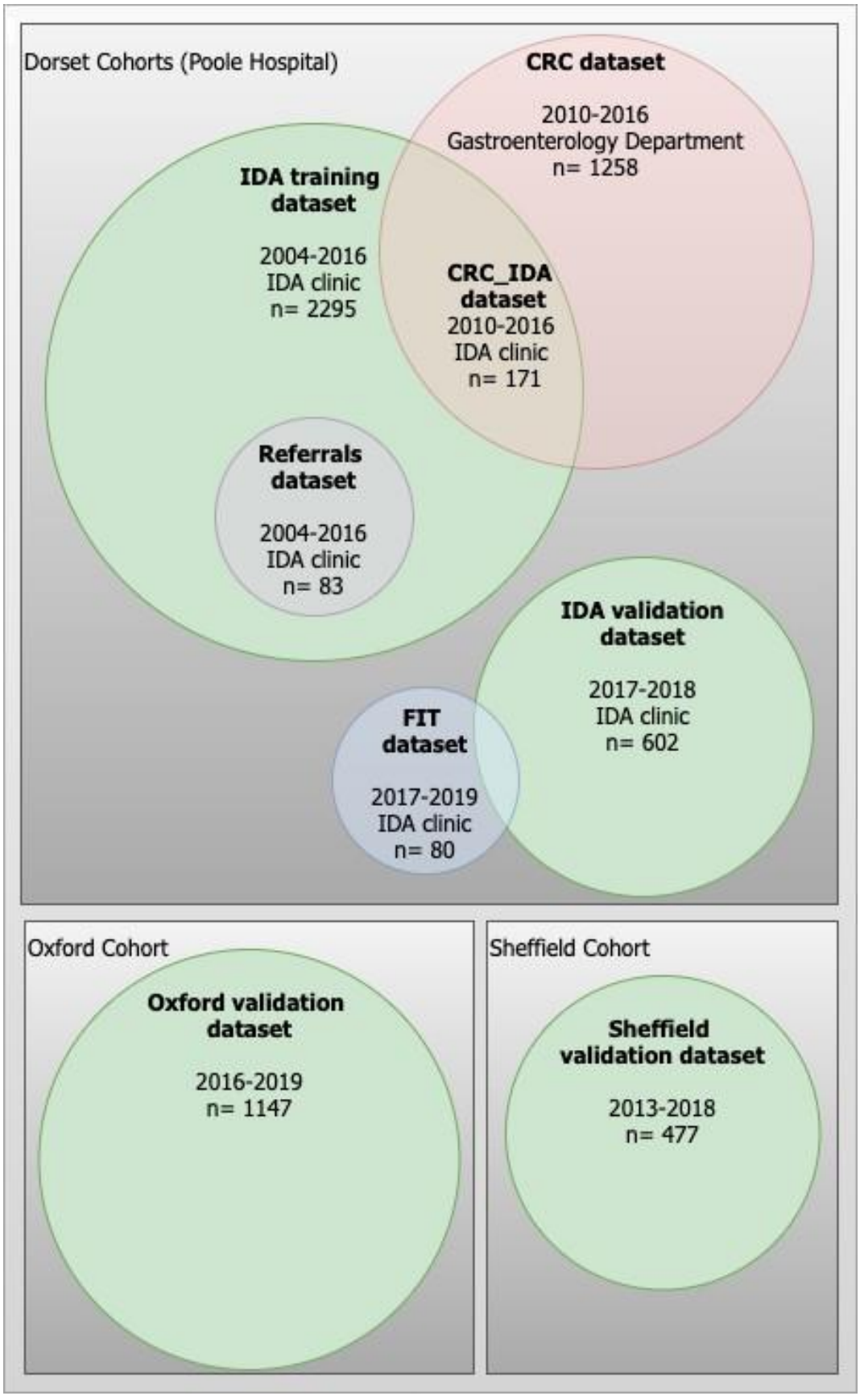


Figure 3-1 Venn diagram of the datasets analysed in this PhD study

3.2.2 Software development project

The project revolves around developing an automated decision-support tool in which anonymised, individual, patient data is entered and GI cancer risk is calculated based on the IDIOM model and displayed immediately. Human-centred design was employed to develop the solution, focusing on the users and their needs, whilst ensuring that they are provided with sufficient details to appropriately interpret the risk score. After a self-assessment for conformity certification was conducted to apply the CE marking, the technical documentation of the *IDIOM App*, is established per Annex VII in the MDD to include: the app general description and its intended use(s); development planning; requirements analysis; implementation, and deployment; clinical evaluation and interface usability assessment; risk analysis, maintenance and plan for post-market surveillance; and release and label. Also, the code, the data, the signed declaration of conformity, and a list of all the harmonization legislations and standards that has been adhered to during the app development and the writing of the technical documentation were included in the technical documentation.

3.3 Main statistical analysis

Different statistical approaches were employed in the PhD project, with each method driven by the research question and the data at hand. Details about the sample size, missing data, outcomes, predictors, etc. are embedded within each individual study, but the following are the main methods used to address different objectives:

1. Multiple logistic models were employed to study the association between age, sex, Hb, and MCV, with GI cancer as the outcome, using *IDA training dataset* (objective 1). And to check whether FIT can improve risk stratification using *FIT dataset* (objective 3).
2. The predictive performance of the original model was evaluated by estimating measures of calibration, discrimination, and clinical utility

using the *IDA validation dataset*, *Oxford validation dataset*, and *Sheffield validation dataset* (objective 2 and 4).

3. The cut-offs selected to create the risk groups were based on the quarters of the positive predictive value (PPV) in *IDA training and validation datasets* combined- in which PPV is defined as the number of positive cases that were correctly classified divided by the total number of positive cases predicted-. The lowest quarter divided into two categories. At the highest cut-off, negative predictive value(NPV) remains 100% (objective 5).

4. Continuous-time multi-state Markov chain was run to determine the transition rates among three observed states for IDA patients at the IDA clinic, “incomplete investigations,” “negative GI cancer,” and “positive GI cancer” and to estimate the delay time using *Referrals Dataset* (objective 6).

5. Logistic regression models were employed to assess the associations between diagnostic pathway (Bowel Cancer Screening Programme (BCSP), IDA, symptomatic) and tumour side/stage using *CRC dataset* (objective 8).

6. Bayesian parametric survival model was employed to examine the relationship between CRC stage and the event of having a blood count prior to CRC diagnosis using *CRC_IDA detailed dataset* (objective 9).

3.4 Data management

The data were collected and anonymised by artificial identifiers by the clinical team in each hospital prior to the data provision and analysis. The PhD researcher had no interaction with patients, and no access to patients' records or to the identifiers mapping code. Before carrying out the analysis on the provided data, the data was cleaned and prepared. The preparation process included detecting any odd values, inconsistency between the variables, duplicates, and human entry errors. Records of cleaning the data, and the data itself are included in appendix II. Any decision to correct the data was discussed, checked, and approved by the clinical and supervisory team before implementation.

The PhD study documents, code and results were stored on the university encrypted H drive into one folder called PhD drive. PhD drive consists of several folders; each was dedicated for specific purposes such as reference, R code, PhD ethical approval documents, etc. The PhD study was conducted in accordance with the Good Research Practice, and ICH GCP, BU Code of Practice, and in accordance with the recommendations for physicians involved in research on human subjects adopted by the 18th world medical assembly, Helsinki 1964 and later revisions.

3.5 Ethical approvals

The studies of objective numbers 1, 2, 4, 5, 6, and 8, were low-risk studies revolve around conducting a retrospective statistical analysis on an anonymised secondary data. The timeline of attaining all the ethical approvals for them:

- Attained the "No need for the NHS approval" authoritative decision on 7th Nov 2017 (appendix IIIa).
- BU ethics checklist approval decision on the 22nd Feb 2018 with Ethics ID: 19925 (appendix IIIb).

- Confirmation of Capacity and Capability at Poole Hospital NHS Foundation Trust on the 3rd May 2018 (appendix IIIc).
- Completing the study BU electronic folder at the BU ethics research as required by Good Clinical Practice guidelines on the 8th May 2018.

- For the pilot study (an observational study following a simple universal intervention (FIT assessment) of objective 3, the ethical approval was attained from the HRA (appendix IIId).

- For the study of objective 7, the participants in the App interface usability assessment questionnaire were NHS staff such as IDA nurse specialists and gastroenterologists, and participation was voluntary and anonymous, so no ethical approval was needed for these participants. The MHRA were consulted on 13th Nov 2018 with regard to the CE marking and risk class. They advised that IRAS application would be required in the case of conducting clinical investigation which was not the case for the *IDIOM* (appendix IIIe).

Chapter 4 : Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia

Open access original research article published at: https://bmjopengastro.bmj.com/content/7/1/e000403

Citation: 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. 2020. *BMJ Open Gastroenterology;* 7:e000403. doi: 10.1136/bmjgast-2020-000403

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 (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 – 1.07, P<0.00001), 2.86 for males (2.03 – 4.06, P<0.00001), and 1.03 for each g/l reduction in Hb (1.01 – 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 – 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 - 51.8), but the sensitivity was low at 23.5% (95% CI 6.8 – 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.

Summary Box

What is already known about this subject? GI cancer is the cause of IDA in 8 – 10% of adult males and post-menopausal females 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 MCV 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

Introduction

Iron deficiency anaemia (IDA) is a common clinical problem, with an overall incidence in western populations approaching two cases per 1,000 pa, and a considerably higher age-specific incidence in those over the age of 70 [1 2]. More than a quarter of males and post-menopausal females 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 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 has 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 co-morbidities. 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 pre-investigation 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 - appear to be independent predictors of underlying GI cancer in IDA. In the IDIOM (Iron Deficiency as an Indicator Of Malignancy) 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 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 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.

Methods

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, whilst 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 (AIC), Anova chi square test, Pseudo R squared 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 appropriate [14] 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 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

A pilot study to explore the potential role of FIT - IDIOM-3 (ISRCTN No 18342140) - was undertaken with Research Ethics approval (IRAS No 201759). 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], (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 Hema-screen 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 (version 3.6.1), RStudio (version 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 females and 862 males (a sex ratio of 1.8), with a median age of 71 years (inter-quartile range: 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 v 106 g/l, P <0.001) and mean MCV (79.4 v 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 confidence interval. 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: $\ln(GI\_cancer) \sim \beta_0 + \beta_1 age + \beta_2 sex + \beta_3 Hb + \beta_4 MCV$. 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 odds ratios (95% CI, P value) for the four predictive variables were as follows:

- **Age** - 1.05 per year (1.03 – 1.07, P<0.00001)
- **Sex** - 2.86 for males (2.03 – 4.06, P<0.00001)
- **Hb** - 1.03 for each g/l reduction (1.01 – 1.04, P<0.0001)
- **MCV** - 1.03 for each fl reduction (1.01 – 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 4-1). Using the 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 sub-group. Cut-off 2 (risk 7.4%) maximised G mean in the training dataset (69.2% ; 95% CI 21.8 – 219.9%) and gave a comparable value in the validation dataset (73.2% ; 95% CI 27.4 – 195.6%), with closely overlapping confidence intervals and similar values for sensitivity and specificity. Cut-off 3 (risk 11.1%) stratified about 25% of both cohorts into a high risk sub-group. These results (summarised in Table 4-1) 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.

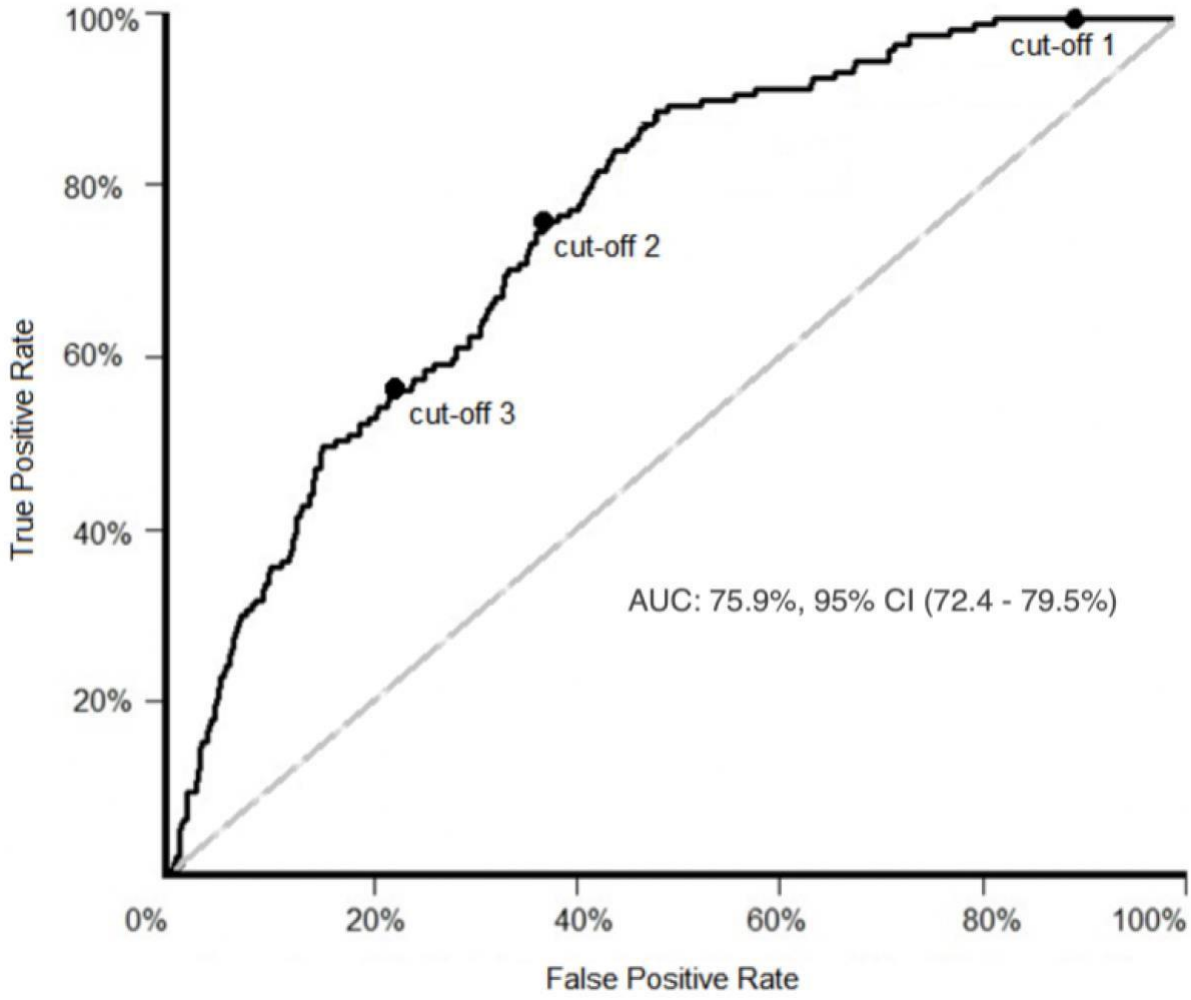


Figure 4-1 Receiver operating characteristic (ROC) curve for the training dataset, showing the three optimal cut-off points defined in the text: cut-off 1 = 1.5%, cut-off 2 = 7.4%, cut-off 3= 11.1%. AUC = area under curve

Table 4-1 Characteristics of the three optimal cut-off points for predicted probability of GI cancer, as applied to the training and validation datasets. NPV = negative predictive value, G mean = geometric mean, PPV = positive predictive value

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

The striking effect of combining the predictive variables on predicted risk is displayed in heat-map format in Figure 4-2. This demonstrates the high risk in all older males with IDA regardless of haematology findings, and the extremely low risk in younger females 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.

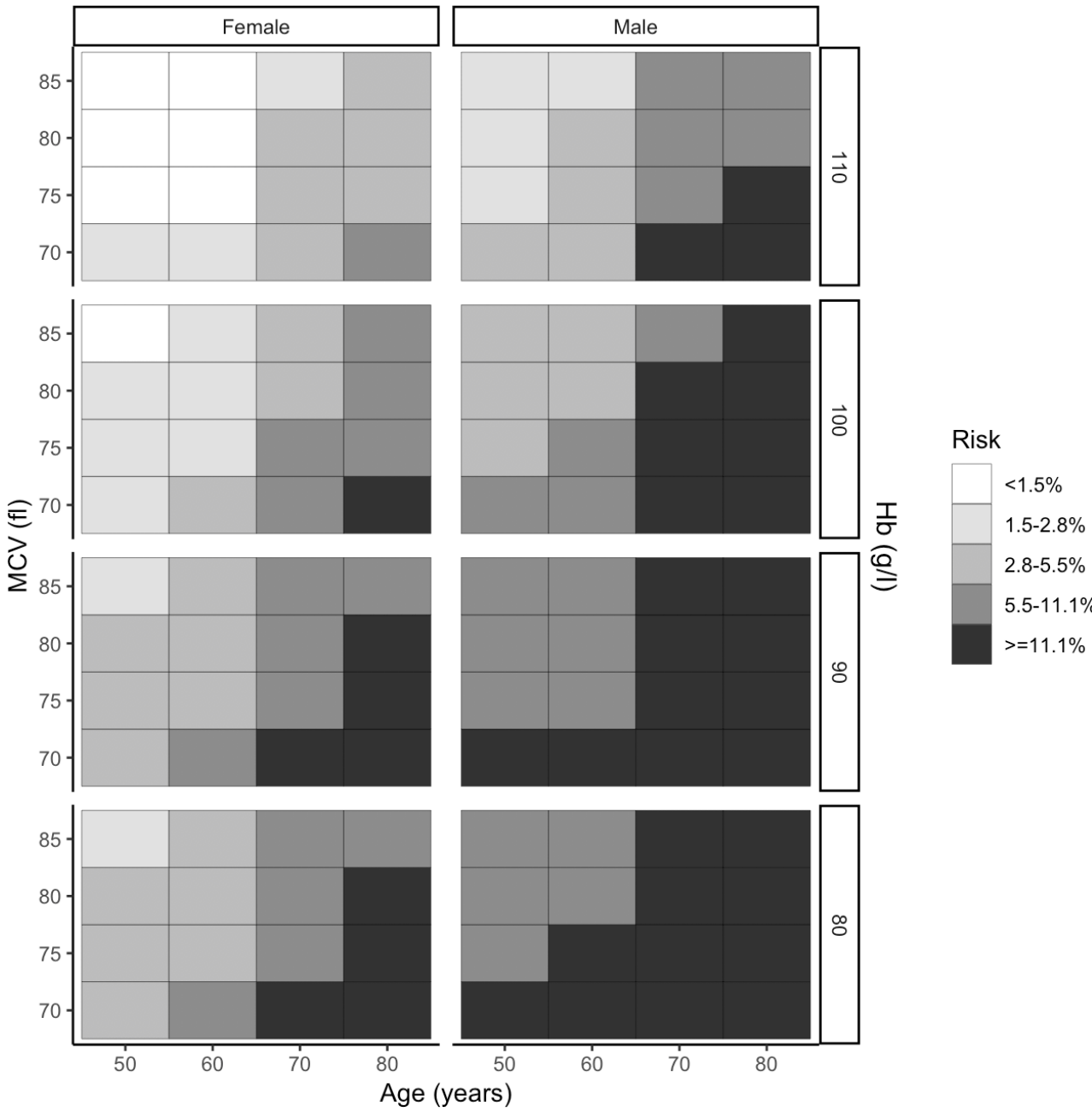


Figure 4-2 Heatmap showing the probability of gastrointestinal (GI) cancer in the overall IDA cohort (n=2390) according to age, sex, blood haemoglobin concentration (Hb: g/l) and mean cell volume (MCV: fl). The darker the box, the higher the GI cancer risk – as shown on the risk key. The risk ranges are based on positive predictive value quartiles, with the lowest quartile divided in two

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 4-2 - FIT positivity was associated with GI malignancy (OR = 6.6, 95% CI 1.6 - 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 – 49.9%), and this only increased to 26.7% (95% CI 7.8 – 55.1%) with exclusion of the upper GI cancers.

Table 4-2 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

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% confidence interval) 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 MHRA approval and CE marking. A screenshot from the App is shown in Figure 4-3.

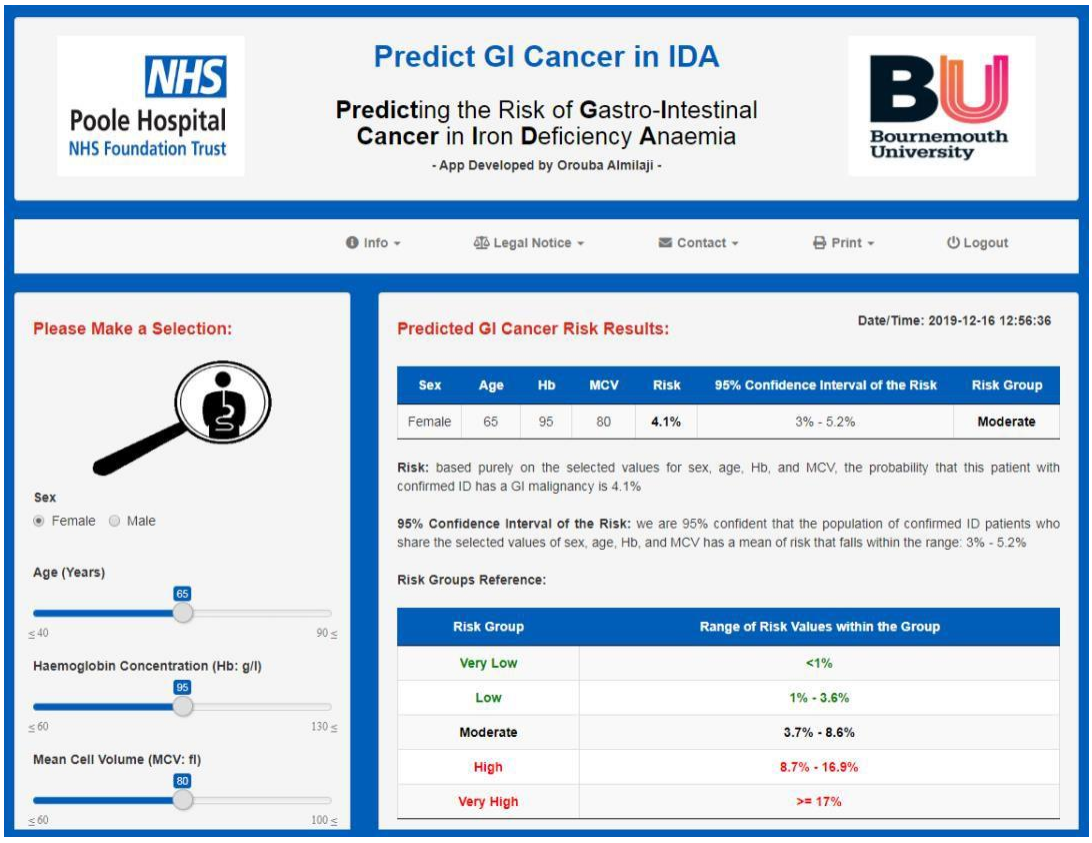


Figure 4-3 A screenshot from the App Predict GI Cancer in IDA

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 co-morbidities, 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 pre-test 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. Firstly, it allows the identification of a high-risk sub-group who warrant accelerated investigation and can be advised accordingly. Secondly 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. Firstly, the predicted GI cancer risk is in all cases greater than 0% and less than 50%. Secondly, whilst 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, whilst 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 a relationship with MCV, albeit less strong – barring a single report on a very small cohort [21] this has not been evident in previous analyses [9 10], and has perhaps emerged in this study because of the substantially larger cohort size. The observations are further strengthened by the findings of the internal prospective validation exercise reported here.

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 males [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]. But 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 whilst 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 sub-groups 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 post-menopausal females, as shown in Figure 4-2. The finding is unlikely to be the result of referral bias, as younger women with mild anaemia are the IDA sub-group 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 commonest GI cancer underlying IDA) than guiac-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 sub-group 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 whilst right-sided lesions account for about 35% of all CRCs, the figure is over 80% for the sub-group 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 5 studies with a sensitivity of 0.82 (95% CI 0.68–0.90), 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.

Statements

Contributorship

OAM, AC, PT, EJW and JAS conceived and designed this study, and CS and SLS collected the data. OAM analysed the data and drafted the initial manuscript, and JAS is the guarantor. All authors made significant contributions to the subsequent revision of the paper, and approved the final version prior to submission.

Funding sources

(1) FIT kits funded by Poole Hospital Gastroenterology Research Fund. (2) PhD studentship (OAM) jointly funded by Poole Hospital Gastroenterology Research Fund and Bournemouth University. (3) AC part-funded by the National Institute for Health Research (NIHR) Applied Health Research Collaboration (ARC NWC). The views expressed are those of the authors and not necessarily those of the NIHR or Department of Health and Social Care.

Competing interests

SLS and EJW have received honoraria for speaking at educational meetings sponsored by Pharmacosmos.

References

1. Camaschella C. Iron deficiency anaemia. New Engl J Med 2015;372:1832-43
2. Stone H, Almilaji O, John C, et al. The dedicated iron deficiency anaemia clinic – a fifteen-year experience [in preparation]
3. Rockey DC. Gastrointestinal tract evaluation in patients with iron deficiency anemia. Semin in Gastrointest Dis 1999;10:53-64
4. Liu K, Kaffes AJ. Iron deficiency anaemia: a review of diagnosis, investigation and management. Euro 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 CM, 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 (second ed)
12. Ferri C, Hernández-Orallo J, Modroiu R. An experimental comparison of performance measures for classification. Pattern Recognit Lett 2009;30:27-38

13. Chen JJ, Tsai CA, Moon H, et al. Decision threshold adjustment in class prediction. SAR QSAR Environ Res 2006;17:337-52
14. Afonja T (2017-12-08). Accuracy Paradox. Towards Data Science [retrieved 15.3.19]
15. Tharwat A. Classification assessment methods. Applied Computing and Informatics 2018. doi: 10.1016/j.aci.2018.88.003
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 doi: 10.1007/3-540-62858-4_79
17. Yu H, Hong S, Yang X, Ni J, et al. Recognition of multiple imbalanced cancer types based on DNA microarray data using ensemble classifiers. Biomed Res Int 2013:239628 doi: 10.1155/2013/239628
18. Lin W-J, Chen JJ. Class-imbalanced classifiers for high-dimensional data. Brief Bioinform 2012;14:13-26
19. Fernández A, García S, Galar M, et al. Learning from Imbalanced Data Sets. Switzerland:Springer 2018 (first ed)
20. Alpha Laboratories UK ; hema-screen SPECIFIC FIT, 2018. Available at: https://www.alphalabs.co.uk/media/productfile/file/h/e/hemascreen-specific-2012.pdf (Accessed: 15.12.19)
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 at: http://www.cancerresearchuk.org/cancer-info/cancerstats/incidence/uk-cancer-incidence-statistics (Accessed: 15.12.19)
24. Burton JK, Yates LC, Whyte L, et al. New horizons in iron deficiency anaemia in older adults. Age Ageing, afz199. https://doi.org/10.1093/ageing/afz199

25. Almond S, Mant D, Thompson M. Diagnostic safety-netting. Br J Gen Pract 2009;59:872-4. https://doi.org/10.3399/bjgp09X472971
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 [published online ahead of print, 2020 Jan 10]. Am J Gastroenterol. 2020;10.14309/ajg.0000000000000495. doi:10.14309/ajg.0000000000000495
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(1):e000293. Published 2019 May 4. doi:10.1136/bmjgast-2019-000293
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 2018 DOI: 10.1002/bjs5.50131
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. Published 2017 Dec 4. doi:10.2147/RMHP.S109116
31. Smith A, Young GP, Cole SR, et al. Comparison of a brush-sampling fecal 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. Gastroenterol 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 2011;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. doi:10.1056/NEJMoa1909159
36. Kim NH, Yang HJ, Park SK, et al. Does low threshold value use improve proximal neoplasia detection by fecal immunochemical test? Dig Dis Sci 2016;61:2685-93. doi: 10.1007/s10620-016-4169-3.

Chapter 5 : Broad external validation of a multivariable risk prediction model for gastrointestinal malignancy in iron deficiency anaemia

Open access original research article published at: https://diagnprognres.biomedcentral.com/articles/10.1186/s41512-021-00112-8

Citation: Almilaji, O., Webb, G., Maynard, A. et al. Broad external validation of a multivariable risk prediction model for gastrointestinal malignancy in iron deficiency anaemia. 2021. *Diagnostic and Prognostic Research;* 5:23. doi: 10.1186/s41512-021-00112-8

Abstract

Background

Using two large datasets from Dorset, we previously reported an internally validated multivariable risk model for predicting the risk of GI malignancy in IDA – the IDIOM score. The aim of this retrospective observational study was to validate the IDIOM model using two independent external datasets.

Methods

The external validation datasets were collected, in a secondary care setting, by different investigators from cohorts in Oxford and Sheffield derived under different circumstances, comprising 1117 and 474 patients with confirmed IDA respectively. The data were anonymised prior to analysis. The predictive performance of the original model was evaluated by estimating measures of calibration, discrimination, and clinical utility using the validation datasets.

Results

The discrimination of the original model using the external validation data was 70% (95% CI: 65, 75) for the Oxford dataset and 70% (95% CI: 61, 79) for the Sheffield dataset. The analysis of mean, weak, flexible, and across the risk groups calibration; showed no tendency for under or over-estimated risks in the combined validation data. Decision curve analysis demonstrated the clinical value of the IDIOM model with a net benefit that is higher than “investigate all” and “investigate no-one” strategies up to a threshold of 18% in the combined validation data. Using a risk cut-off of around 1.2% to categorise patients into the very-low risk group showed that none of the patients stratified in this risk group proved to have GI cancer on investigation in the validation datasets.

Conclusion

This external validation exercise has shown promising results for the IDIOM model in predicting the risk of underlying GI malignancy in independent IDA datasets collected in different clinical settings.

Keywords: Iron deficiency anaemia; gastrointestinal cancer; IDIOM app; external validation; temporal validation; TRIPOD.

Background

The strong association between iron deficiency anaemia (IDA) and gastrointestinal (GI) cancer is well recognised [1-5]. As a result, IDA in at-risk groups is an accepted indication for fast-track referral to secondary care for further investigation in the UK [6]. The problem with this approach is that IDA is common, but the prevalence of malignancy is only 8 - 10% [7] - meaning a large workload for a relatively small yield.

With the aim of risk stratification, we have previously built and internally validated a binary multivariable logistic model to predict the risk of GI cancer in patients with confirmed IDA, based on four simple variables: age, sex, haemoglobin concentration (Hb), and mean cell volume (MCV) – the IDIOM model (Iron Deficiency as an Indicator of Malignancy) [4]. Identifying subgroups of IDA patients who are at increased / reduced risk of GI cancer might lead to (a) accelerating the investigation of those at high predicted risk, with potential prognostic implications, and (b) helping those at low predicted risk to avoid unnecessary invasive procedures.

The clinical data used to develop the original model was collected for adult patients with IDA (n=1879) referred to the IDA clinic in Poole hospital during the period 2004-2016 inclusive. The criteria for inclusion were iron deficiency confirmed by standard laboratory criteria, and subsequent investigation of the upper and lower GI tract. Due to informed patient preference, concurrent illness, or major co-morbidity, about 10% of IDA patients usually fail to undergo GI investigation for IDA [3,8].

Developing the original model was carried out before receiving an internal dataset (n=511) from the later period 2017-2018. Using this dataset, temporal validation of the fitted model showed excellent promise of generalisability. The

area under the receiver operating characteristic (ROC) curve (AUC) was estimated at 81% (95% CI: 74, 86). The prevalence of malignancy was 8.4% in the temporal dataset. The average estimated risk of 7.6% indicated that the IDIOM model has no tendency to underestimate or overestimate risk. The calibration intercept and slope were 0.1 (95% CI: -0.1, 0.4) and 1.1 (95% CI: 0.8, 1.5) respectively, suggesting that risk estimates were not systematically too moderate or extreme.

However, the data used to validate the model were collected by the same centre (IDA clinic in Poole hospital), for the same population (Dorset), using the same predictors and outcome definitions and measurements. Confirmed IDA was defined using the same blood laboratory marker cut-offs in the training and internal validation datasets, but these cut-offs are relevant only to the local laboratory and may vary between laboratories.

So to apply the model with confidence to different populations it must be tested, and amended in case of poor performance, using data collected by other investigators in other geographic areas and preferably using locality-specific definitions for the predictors. This retrospective cohort study aims to address the transportability of IDIOM score model by broadly validating it using two independent external datasets.

Methods

After temporally validating the model in 2020, the training and internal datasets were merged to form the Dorset dataset. This was used to fit the full IDIOM model [9]. The multiple binary logistic regression of this full model was constructed according to the formula:

$$\log\left\{\frac{\mathbb{P}(GI\ Maligancy = postive)}{\mathbb{P}(GI\ Maligancy = negative)}\right\} = -1.84 + 0.94\ sex + 0.06\ age - 0.03\ MCV - 0.03\ Hb$$

The full IDIOM model was almost identical to the original model using only the training dataset [4]. Statistical assessment of the validity and goodness of fit of the logistic regression model (Smoothed scatter plot, deviance and residual test, Cook's distance and standardised residual errors, variance inflation factor, Akaike information criterion, analysis of variance χ2 test, pseudo R^2) was satisfactory.

Before importing the coefficients of the full IDIOM model to predict the risk of GI cancer in the validation data, least absolute shrinkage and selection operator (Lasso) was applied to regulate the model. A comparison of different regularisation method effects on the model coefficients is shown in the Supplementary information, Table S 5-1. The model coefficients after applying these methods were very close, however, Lasso method was selected because it is the method that shrunk the coefficients the most.

The final updated multiple binary logistic regression of the full IDIOM model regulated using Lasso method and validated in this study was constructed according to the formula:

$$\log\left\{\frac{\mathbb{P}(GI\ Maligancy = postive)}{\mathbb{P}(GI\ Maligancy = negative)}\right\} = -1.84 + 0.89\ sex + 0.05\ age - 0.03\ MCV - 0.06\ Hb$$

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

- Sex: 2.44 for men (1.88 to 3.49, p<0.0001)
- Age: 1.05 per year (1.04 to 1.08, p<0.0001)
- MCV: 1.03 for each fl reduction (1.01 to 1.05, p<0.01)
- Hb: 1.03 for each g/l reduction (1.02 to 1.04, p<0.0001)

The quartiles of positive predictive values (PPV) were updated based on the penalised model (Table 5-1). The first PPV quarter was divided into two halves, in which the lower half corresponds to negative predictive values (NPV) equal to 100% only. The highest predicted risks in each PPV quarter

(and the lower half of the first quarter) were used as cut-offs to create the risk groups. The updated cut-offs to create the risk groups were 1.18%, 2.16%, 4.24%, and 7.97%.

Table 5-1 Risk groups cut-offs after regulating IDIOM model based on the quartiles of PPV

PPV quarters	PPV values range %	Corresponding predicted risk cut-offs %	Risk group
Lower half of the 1st quarter of PPV*	[8.4-9.4]	<=1.18	very-low risk
Upper half of the 1st quarter of PPV	]9.4-10.8]	]1.18-2.16]	Low risk
2nd quarter of PPV	]10.8-14.7]	]2.16-4.24]	Moderate risk
3rd quarter of PPV	]14.7-19.6]	]4.24-7.97]	High risk
4th quarter of PPV	>19.6	>7.97	Very high risk

*: The risk group at which PPV values are in the lower quarter, and NPV =100

PPV is the number of positive cases that were correctly classified divided by the total number of positive cases predicted. NPV is the number of negative cases that were correctly classified divided by the total number of negative cases predicted.

The highest Gmean (geometric mean of sensitivity and specificity) value in the Dorset dataset was updated using the penalised IDIOM model and found to be around 70%.

Source of data

Independent datasets were collected by investigators in Oxford and Sheffield and included all subjects who met the inclusion criteria within the collection time frame. The Oxford dataset was collected for the period 2016-2019 and comprised 1147 subjects with IDA referred for fast-track investigation. The Sheffield dataset was collected for the period 2013-2018 and compromised 477 subjects with IDA referred to a dedicated IDA Clinic.

Patients

For all datasets the subjects were adults referred to secondary care who went on to be investigated for IDA. The decision to refer was generally made in primary care by the GP who requested the blood test revealing IDA, usually following a discussion with the patient concerned about the significance and potential implications of the result.

Confirmation of IDA depended on local practice but was broadly accepted as: (a) Transferrin saturation (T.sat) <15% and / or serum ferritin less than the lower laboratory limit of normal at the time for the Dorset dataset, (b) T.sat <16% and / or serum ferritin <10 µg/l (women) or <20 µg/l (men) for the Oxford dataset, (c) serum ferritin <31 µg/l (both sexes) for the Sheffield dataset. The diagnosis of iron deficiency was confirmed in all subject in each of the datasets by the finding of an abnormally low serum ferritin and / or transferrin saturation. All subjects underwent standard first-line GI investigation for IDA, comprising exclusion of coeliac disease, OGD and an adequate colonic examination – either by CT colonography or colonoscopy.

Outcome and predictors

As for the IDIOM model, the outcome was the presence / absence of cancer of the upper or lower GI tract. The predictors were the recoded values of age at presentation (years), sex (male / female), Hb (g/l) and MCV (fl) measured from the same blood sample taken prior to iron replacement therapy. The decision regarding the presence or absence of GI malignancy was made by clinician with responsibility for the case after GI investigations were complete.

Sample size

Being a retrospective analysis of secondary data meant there was no control of the size of the external validation datasets. The number of outcome events in the Oxford and Sheffield datasets were 86, 36 respectively. Following the simulation-based sample size calculations for external validation of clinical prediction models [10], the anticipated precisions of performance measures were estimated based on the available number of outcome events in each external validation dataset, and on them both combined. Further details about sample size considerations are included in the Supplementary information (sample size, Figure S 5-1, Figure S 5-2).

Development vs. validation(s)

Differences in the quoted normal ranges for Hb, MCV, T.sat, and serum ferritin were to be expected between the laboratories in Dorset, Oxford, and Sheffield as these references are relevant only to the local laboratory. However, the differences for all the variables (as shown in the Supplementary information, Table S 5-2) were marginal.

Missing data

There was no missing data in the external validation datasets for the results of IDA investigations, Hb, MCV, sex, and age.

Statistical analysis

Before starting the analysis, external validation datasets were prepared by taking out duplicates and applying inclusion (confirmed IDA patients who underwent standard first-line GI investigation for IDA) and exclusion criteria (all IDA patients diagnosed with other malignancies e.g. ovarian cancers, renal cancers, GIST, neuroendocrine tumours were excluded). The updated cut-offs used to create the risk groups in Dorset (Table 5-1) were imported to create risk groups in Oxford, Sheffield, and the combined validation datasets, and then the predictive performance of the IDIOM model was evaluated using the validation datasets by estimating the following measures:

Discrimination

Discrimination refers to the ability of the model to distinguish correctly between the presence and absence of GI cancer in the validation datasets. Discrimination of the IDIOM model was assessed by examining the values of C-statistic for these datasets. For a binary outcome, the C-statistic is equivalent to the area under the receiver operating characteristic (ROC) curve (AUC). The highest Gmean values in the Dorset, Oxford, Sheffield, and the combined validation datasets were compared visually by adding them to the ROC curve graph.

Calibration

Calibration quantifies how close estimated risks are to observed ones in the validation datasets. Assessment of IDIOM calibration was carried out following published methodology [11], employing mean calibration (or calibration-in-the-large), weak calibration (calibration intercept and calibration slope), and moderate calibration (flexible calibration curve based on Loess functions).

To check calibration across the risk groups, we split the combined validation dataset based on descending order of probabilities into fifths (5 groups) using the defined cut-offs (in Table 5-1). Then the calibration between observed and predicted risks across the risk groups was assessed visually using a calibration plot. As per the sample size considerations (Supplementary information), the two external datasets were combined to assess the calibration.

Net benefit

The net benefit (NB) expresses the relative value of benefits and harms associated with using the model. Benefits reflect the diagnosis of a GI cancer by investigation, whilst harms include the risks and cost of carrying out an unnecessary invasive investigation.

Since the current standard of care is to offer investigation to all patients with IDA at risk of malignancy, a major potential use of the IDIOM risk model would be to identify those at very low risk who may not warrant investigation.

Decision curves can be used as a tool for assessing the performance of risk prediction models [12], Decision curves were used to assess the clinical value of the model by ensuring that it had a higher NB than simple strategies such as “investigate all” or “investigate no-one” across a plausible range of risk thresholds. Clinical impact curves, which are alternative plots for the outputs of decision curves, were used to compare the estimated number of patients who would be classified as low risk, and the number of patients classified low risk without the outcome of interest (true negative) at each threshold.

Subjects diagnosed with GI cancer (cases) have expected benefit B > 0 from the investigation, where B accounts for the totality of good and bad effects. Likewise, subjects who do not have GI cancer (controls) have a cost (or burden) of the investigation, C > 0 [13]. Given benefit (B) and cost (C), the optimal risk threshold (R) for determining investigation is:

$$R = \frac{C}{C + B}$$

When the policy is “investigate all”, all controls experience the cost of investigation. The advantage of an opt-out policy to the patient population accrues from controls whose estimated risks are below R, as such patients avoid the cost [13]. Expressing NB in terms of avoided unnecessary investigations is recommended if the reference strategy is “investigate all”, and so NB is expressed in terms of true negatives rather than true positives [14]. Given that ρ is the proportion of cases, the standardised net benefit- which is easier to be interpreted than net benefit- can be calculated by dividing the net benefit by $(1 - \rho)$ as can be shown from the equation [13]:

$$sNB = TNR_R - \frac{\rho}{(1-\rho)} \frac{1-R}{R} FNR_R$$

In which TNR is the specificity at a given risk threshold R, and FNR is the miss rate at the same threshold. At “investigate no-one”; TNR=FNR≡1, and at “investigate all”; TNR=FNR≡0. In an “investigate all” standard of care, the standardised net benefit can be viewed as the TNR appropriately discounted by the FNR [13].

With an “investigate all” standard of care it is difficult for any model to perform better than a strategy of “investigate no-one” when the prevalence is low, and so for this analysis we combined both external validation datasets into one and compared that to the Dorset dataset.

The TRIPOD (Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) Initiative was followed to report this study [15]. R (version 3.6.1) and RStudio (version 1.2.5001) were used to run the statistical analysis and to produce the graphs.

Results

Patients

After tidying the databases and applying the exclusion criteria, 1117 cases were available for detailed analysis from the Oxford dataset and 474 from the Sheffield dataset. There were differences between the datasets, as shown in Table 5-2. As expected, the Oxford dataset had a lower median Hb in particular, as subjects presented exclusively through the fast-track pathway.

Table 5-2 Descriptive statistics for the three datasets

Dataset		**Dorset**	**Oxford**	**Sheffield**
Dataset size	**N**	2390	1117	474
GI Cancer	**positive - n (%)**	200 (8.4%)	86 (7.7%)	36 (7.6%)
Sex	**male - n (%)**	862 (36%)	446 (40%)	227 (48%)
Age (Years)	**median (min, max)**	71 (16, 96)	74 (22, 97)	69 (18,93)
Hb (g/l)	**median (min, max)**	104 (32,159)	91 (29,129)	104 (54,152)
MCV (fl)	**median (min, max)**	80 (53,112)	81 (55,125)	80 (32,104)

A Density plot of continuous variables according to the presence/absence of GI cancer in each dataset using the IDIOM model is illustrated in the Supplementary information, Figure S 5-3 whilst the probability distributions for each dataset are shown in the Supplementary information, Figure S 5-4.

Model performance

Discrimination

The discrimination of the IDIOM model was AUC: 77% (95% CI: 74, 80) for the Dorset dataset, AUC: 70% (95% CI: 65, 75) for the Oxford dataset, AUC: 70% (95% CI: 61, 79) for the Sheffield dataset, and AUC: 69% (95% CI: 65, 74) for the combined validation dataset. As predicted by the sample size calculations, due to the small sample size of the Sheffield dataset, the width of the CI for the discrimination was the largest. And for the combined validation data, was less than 10%.

Using the risk groups cut-offs in Table 5-1, analysis showed that:

- Cut-off 1 (<=1.18%) stratified about 11% of the Dorset dataset and 3% of both external cohorts (Oxford:2%, Sheffield: 5%) into the very-low risk group.
- Cut-off 2 (1.18-2.16%) stratified about 14% of the Dorset dataset and 8% of both external cohorts (Oxford:7%, Sheffield: 12%) into the low risk group.
- Cut-off 3 (2.16-4.24%) stratified about 26% of the Dorset dataset and 25% of both external cohorts (Oxford:22%, Sheffield: 31%) into the moderate risk group.
- Cut-off 4 (4.24-7.97%) stratified about 24% of the Dorset dataset and 31% of both external cohorts (Oxford:30%, Sheffield: 33%) into the high risk group.
- Cut-off 5 (>7.97%) stratified about 25% of the Dorset dataset and 33% of both external cohorts (Oxford:39%, Sheffield: 19%) into the very-high risk group.

The proportion of patients who fell into the higher-risk groups from the Oxford dataset was large (69%). This was expected because the patients in the Oxford dataset had lower Hb values. None of the patients stratified in the very-low risk group from the validation datasets (Oxford:2%, Sheffield: 5%) proved to have GI cancer on investigation as NPV remains 100%.

The ROC curve (Figure 5-1) showed that the highest Gmean values in the validation datasets were close (70% in the Dorset dataset, 66% in the Oxford dataset, 68% in the Sheffield dataset, and 64% in the combined dataset).

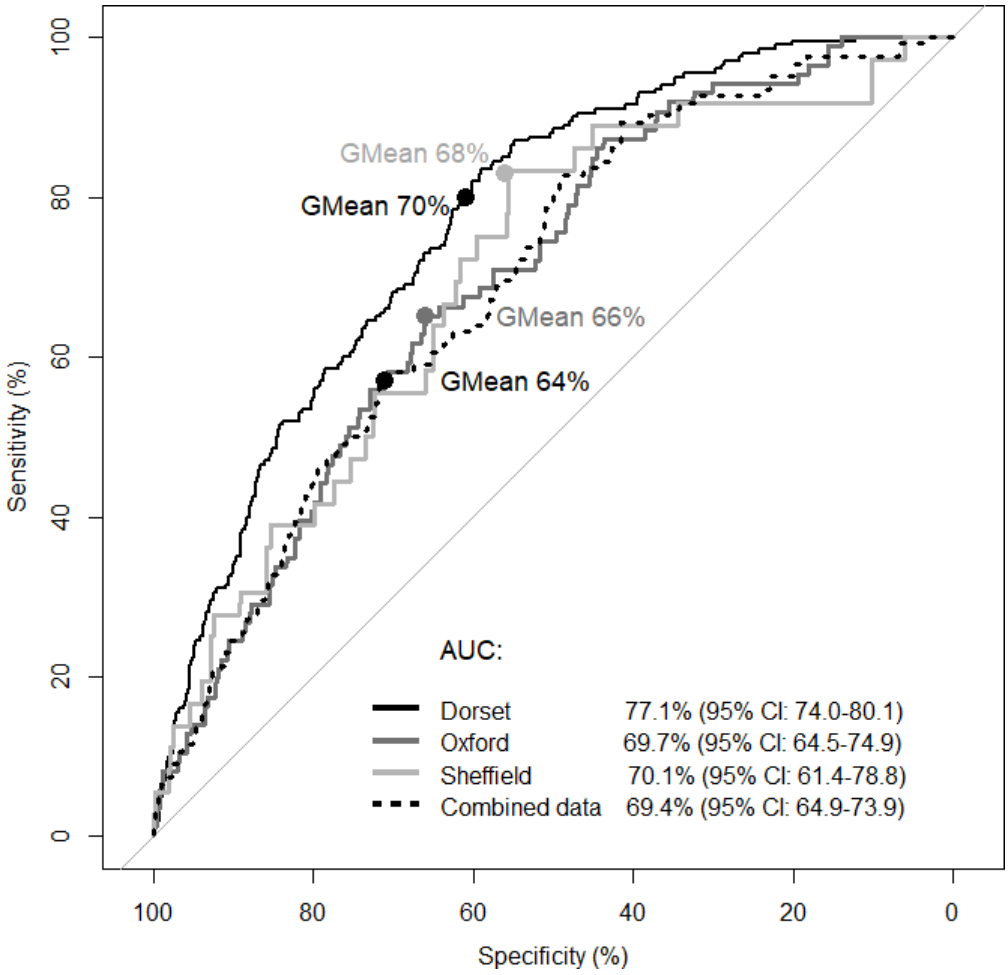


Figure 5-1 Receiver operating characteristic curve shows the sensitivity on Y axis, and specificity on X axis for the Dorset (black), Oxford (dark grey), Sheffield (grey), and combined validation (dotted black) datasets with the highest GMean value in each dataset. AUC - area under curve. GMean - geometric mean of sensitivity and specificity

Calibration

Risk groups calibration

Assessing the calibration visually across the five risk groups in the combined validation data suggested (Supplementary information, Figure S 5-5) that the observed and predicted risks across the five risk groups were overall similar.

Mean calibration (calibration-in-the-large)

The prevalence of malignancy was 7.7% (86/1117) for the Oxford series, 7.6% (36/474) for the Sheffield series, and 7.7% (122/1591) for the combined datasets. The average risks estimated by the IDIOM model were 8.5%, 5.5%, and 7.6% respectively. Assessing the risk ratios, using the validation datasets separately, showed that there was an over-estimating for the risks in the Oxford dataset (by 10%), and under-estimation for the risk in the Sheffield dataset (by 28%). However, the analysis using the combined validation dataset showed no tendency for the model to under- or over-estimate risk (by 1%).

Weak calibration (calibration intercept and calibration slope)

For the Oxford dataset, the calibration intercept and slope were -0.11 (95% CI: -0.34, 0.12), and 0.87 (95% CI: 0.59, 1.15) respectively. For the Sheffield dataset, the calibration intercept and slope were 0.35 (95% CI: 0.01, 0.70) and 0.96 (95% CI: 0.5, 1.42) respectively. For the combined Oxford and Sheffield datasets the number of events was 122, the calibration intercept and slope were 0.01 (95% CI: -0.18, 0.20), and 0.84 (95% CI: 0.60, 1.07) respectively. With zero as the target value for the intercept, the results for Oxford dataset and the combined data suggest no tendency for under- or over-estimated risks. The calibration slopes were close to the target value of 1, suggesting that risk estimates for Oxford dataset and the combined data were not systematically too moderate or extreme in either dataset.

The confidence intervals for the calibration intercept all contain 0 apart from the small size Sheffield dataset. All confidence intervals were wide >0.2, however, this result was consistent with what the sample size calculations predicted based on the existing relatively small number of outcome events.

Moderate calibration (flexible calibration curve)

The flexible calibration plot for the combined validation dataset (Figure 5-2) showed that the model was well-calibrated for risks up to about 17%, but miscalibrated for a few of the higher risk patients. For example, in Figure 5-2, a predicted risk of 30% corresponds to an observed risk of around 20%. However, about 92% of the combined cohort patients have predicted risks less than 17.5% and the model is well-calibrated in this region. Also, using any of these cut-off values above 17.5% would put these patients in the very high-risk group regardless of their predicted risks.

The calibration plot for the Oxford dataset (Supplementary information, Figure S 5-6) showed similar results to that in the combined validation dataset (in which 90% of the Oxford cohort patients have predicted risks less than 17.5%).

Furthermore, the flexible calibration plot for the Sheffield dataset (Supplementary information, Figure S 5-7) showed a strong deviation from the ideal line across the range of true risks above 20%. The miscalibration above 17.5% was consistent with the previous results in the Oxford and combined datasets. However, only 2% of the Sheffield cohort have predicted risks more than 17.5%.

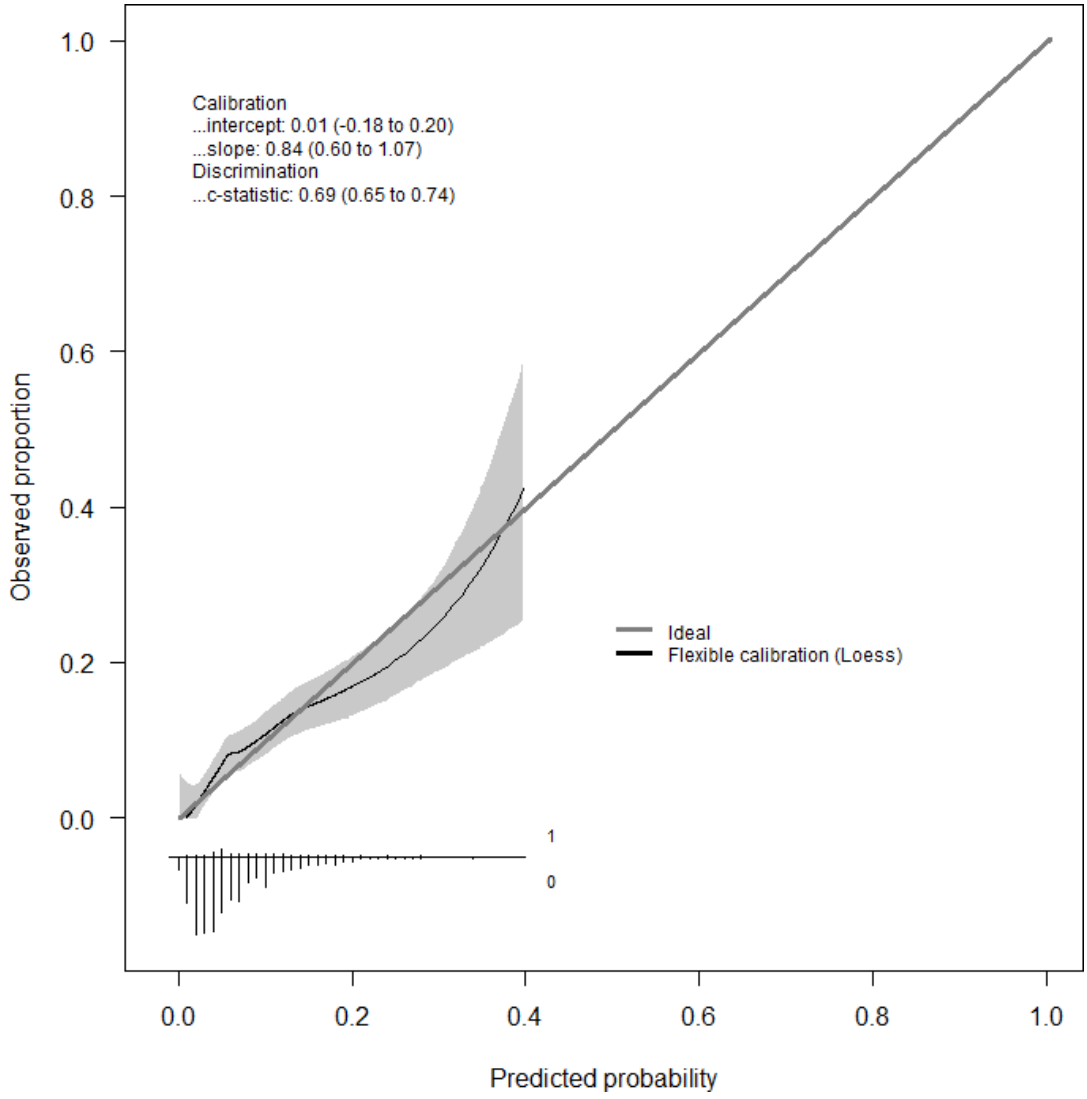


Figure 5-2 Flexible calibration curve for the combined external datasets, showing the relationship between the estimated risks (on the x-axis) and the observed proportion of events (on the y-axis)

Net benefit

Decision curve analysis suggested that the IDIOM model is of clinical value because it has the potential to add value - i.e. standardised NB is higher than "investigate no-one" and "investigate all" for a range of risk thresholds up to 27% in the Dorset dataset, up to 18% in the combined validation dataset (and up to 18% in the Oxford dataset, and to 18% in the Sheffield dataset as can be seen from the Supplementary information, Figure S 5-8 and Figure S 5-9).

So at a risk threshold of 10% for example, use of the IDIOM model would be the equivalent to a theoretical strategy that reduced the number of unnecessary investigations by about 43 per 100 in the Dorset dataset (Figure 5-3), and 38 per 100 in the combined validation dataset (38 per 100 in the Oxford dataset, 37 per 100 in the Sheffield dataset) (Figure 5-4).

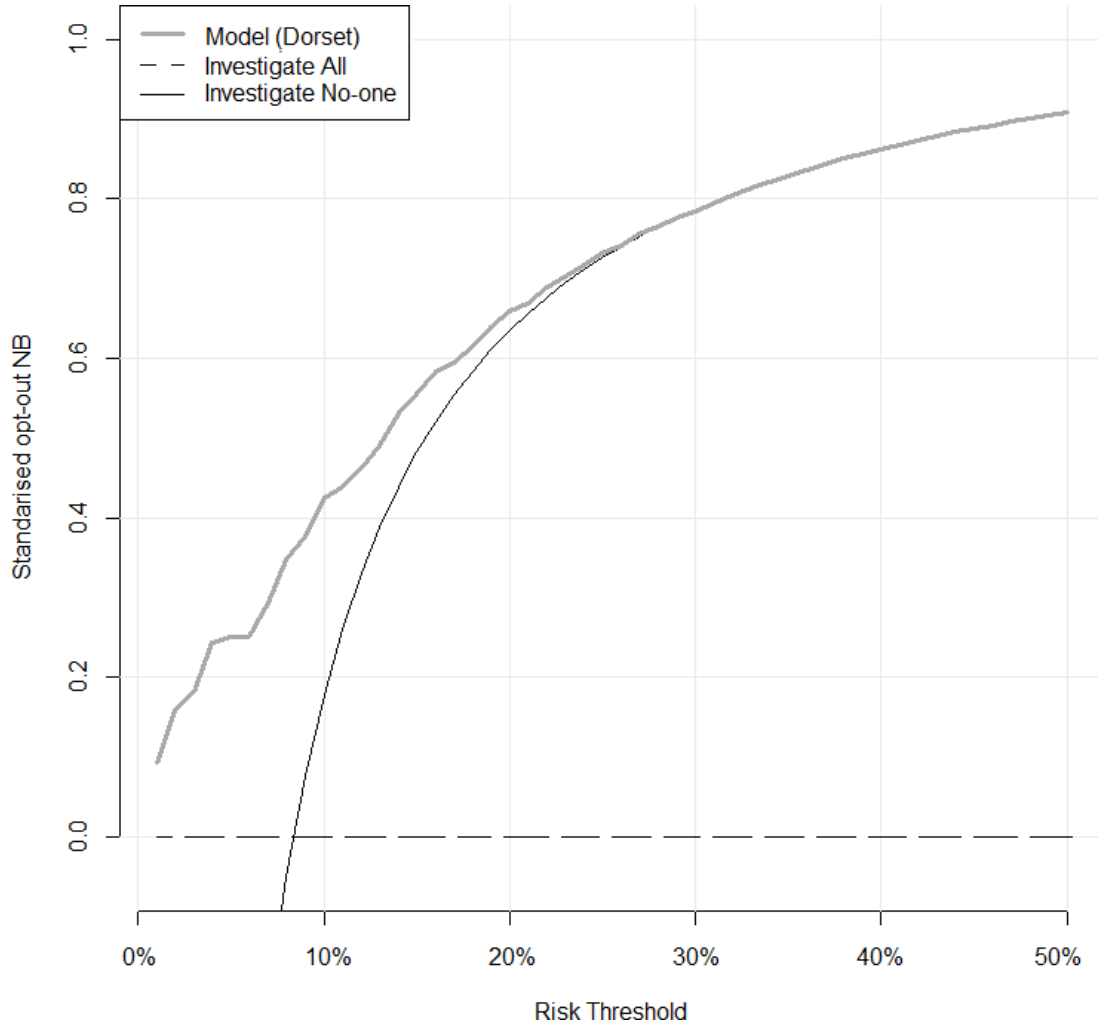


Figure 5-3 Decision curve analysis for GI investigation using Dorset data. Grey line: penalised IDIOM model. Black line: investigate no-one strategy. Dashed line: investigate all strategy. The vertical axis displays standardized net benefit. The horizontal axis shows the risk thresholds

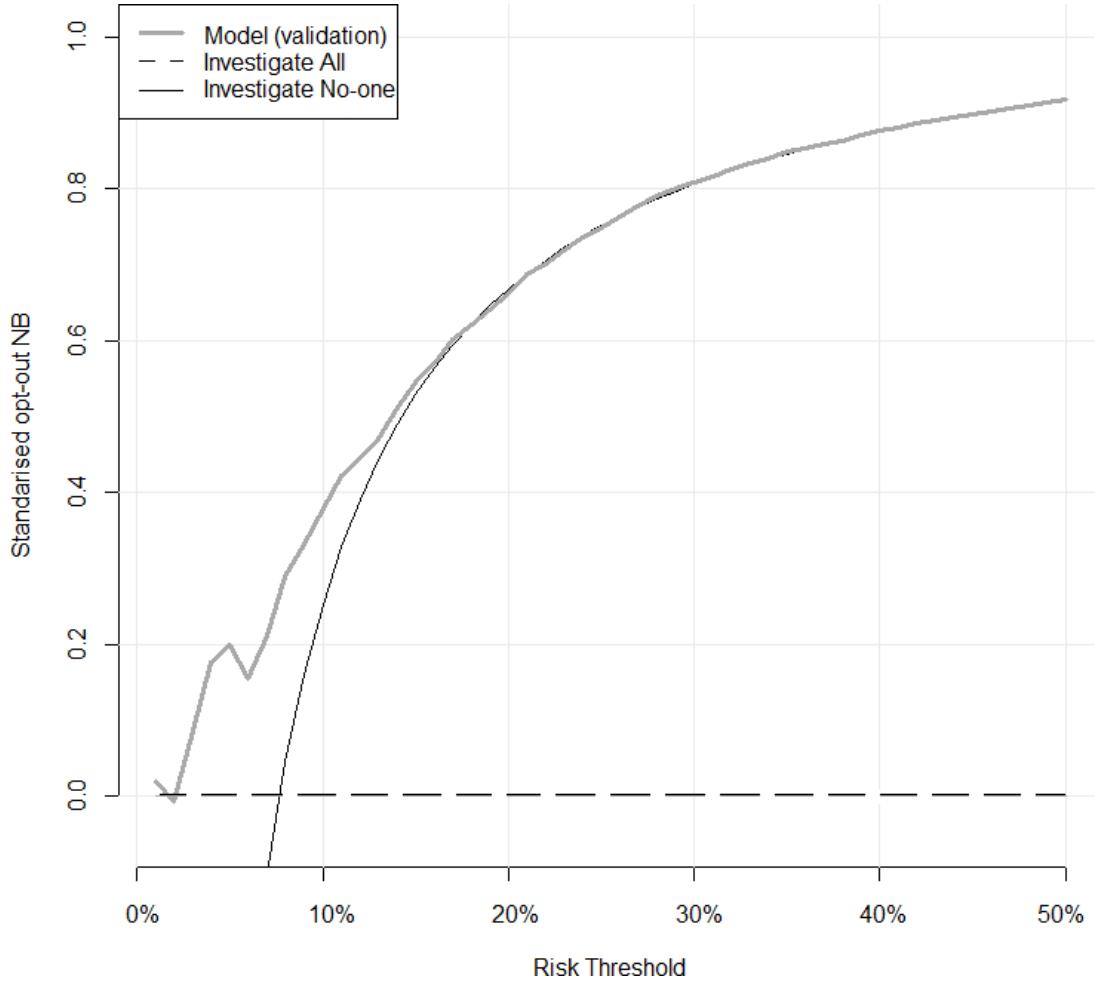


Figure 5-4 Decision curve analysis for GI investigation using the combined external datasets. Grey line: penalised IDIOM model. Black line: "investigate no-one" strategy. Dashed line: "investigate all" strategy. The vertical axis displays standardized net benefit. The horizontal axis shows the risk thresholds

The clinical impact curves (Supplementary information, Figure S 5-10 and Figure S 5-11) showed that at a risk threshold of 10%, around 825 of 1,000 IDA patients in the Dorset dataset would be anticipated as low risk and about 780 of these as true negatives for GI cancer (cost/benefit: 1/9). At the same risk threshold, of 1000 IDA patients in the combined validation dataset, about 750 would be predicted as low risk and about 710 of them as true negatives for GI cancer (cost/benefit: 1/9).

Figure S 5-12 and Figure S 5-13 (Supplementary information) showed that at the same risk threshold of 10%, around 690 of 1,000 IDA patients in the Oxford dataset would be anticipated as low risk and about 650 of these as true negatives for GI cancer (cost/benefit: 1/9). At the same risk threshold, of 1000 IDA patients in the Sheffield validation dataset, about 870 would be predicted as low risk and about 840 of them as true negatives for GI cancer (cost/benefit: 1/9).
Regardless of the sample size, in every dataset, the net benefit analysis has shown consistently a clinical value for the IDIOM model in the validation datasets.

Discussion

IDA is a problem commonly encountered in clinical practice, and the prevalence of underlying GI cancer is the primary justification for the urgent investigation of it [7, 16-21]. Bidirectional endoscopy, 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 [20]. It does however carry a small but significant risk of complications, particularly in the elderly and those with major comorbidities, and it is therefore important to consider the risk-benefit ratio for the investigation of IDA on an individual case basis.

Bidirectional endoscopy 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 of up to 20% of all diagnostic examinations [3]. 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.

We have previously proposed the IDIOM score as a simple and reliable pre-test predictor of the risk of underlying malignancy that is sufficiently discriminating to be clinically useful for patient-centred counselling [4]. Effective risk stratification is a potentially important clinical tool for two reasons. First, it allows the identification of a very high-risk subgroup who warrant accelerated investigation and can be managed 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. Since there is currently no consensus on the risk threshold warranting investigation for GI cancer in IDA, the IDIOM score is of potential use not just to predict the GI risk and stratify patients in meaningful risk groups, but also to inform the decisions of clinicians and patients when discussing whether invasive investigation is appropriate.

Challenges to the applicability of the IDIOM score to other IDA populations include relatively small proportion of positive cases (8.4% for the Dorset dataset), and differences in predictor definitions, referral pathway, and patient characteristics between cohorts in different parts of the country. The external validation exercise reported here was therefore important to confirm that the model underlying the IDIOM score is capable of predicting the risk of underlying GI malignancy in independent external IDA datasets.

Using the combined validation dataset, our results demonstrate that the IDIOM model has good discrimination performance, and of clinical value. The results also suggest that the IDIOM model has no tendency to under- or over-estimate risk, and the risk estimates are not systematically too moderate or extreme. Moreover, using the 1.18% cut-off to categorise patients into the ultra-low risk group showed that none of the IDA patients within this group proved to have GI cancer on investigation in any dataset (Dorset, Sheffield, and Oxford).

The strength of this study is the inclusion of more than one independent external dataset to validate the model. Also, it represents the first risk

prediction model for gastrointestinal cancer in iron deficiency anaemia to be internally and externally validated. Being a retrospective analysis, limitations include our inability to control the size of the study external validation datasets which resulted in a restricted suboptimal evaluation per centre, or to incorporate other variables that might influence GI cancer risk such as family history, previous cancer, race, unintentional weight loss, and red meat consumption - though this is the aim of work to develop the model further.

Conclusion

By analysing two independent datasets, this paper externally validates the IDIOM score risk prediction model, a multivariable logistic regression model developed to predict the risk of gastrointestinal malignancy for patients with iron deficiency anaemia. The assessment of the model performance was evaluated by estimating the measures of discrimination, calibration, and net benefit. This external validation exercise has shown promising results regarding using the IDIOM model in predicting the risk of underlying GI malignancy in different IDA populations in the UK, however, further validation of this model in larger datasets would still be useful to confirm the findings from this study.

Abbreviations

AUC	Area Under The ROC
CI	Confidence Interval
CRC	Colorectal Cancer
GI	Gastrointestinal
Hb	Blood Haemoglobin Concentration
IDA	Iron Deficiency Anaemia
IDIOM	Iron Deficiency as An Indicator of Malignancy
MCV	Mean Cell Volume
NB	Net Benefit
OGD	Gastroscopy
OR	Odds Ratio
ROC	Receiver Operating Characteristic Curve
T.sat	Transferrin Saturation

Declarations

Ethical approval and consent to participate

Retrospective analysis of anonymised secondary data, formal research ethics approval was not required.

Consent for publication

Not required.

Availability of data

Further permission and reasonable re-use requests of the data should be made to the authors in each university.

Competing interest

The authors declare that they have no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Authors' contributors

OAM, EJW, SD and JAS conceived and designed this study. GW and AM provided the Oxford and Sheffield data respectively. OAM analyzed the data and wrote the draft. All authors contributed to the subsequent revision of the paper and approved the final version prior to submission.

Acknowledgment

The authors would like to thank Professor Peter Thomas for his contribution in this study. The authors also would like to thank Professor Richard D Riley for his advice on an early draft of the paper. The authors would like also to thank the two anonymous reviewers for their critical readings, insightful comments, and useful suggestions which led to improving and clarifying this manuscript substantially.

References

1. Hung N, Shen C, Hu Y, et al. Risk of Cancer in Patients with Iron Deficiency Anemia: A Nationwide Population-Based Study. PLoS One 2015;10(3). doi: 10.1371/journal.pone.0119647.
2. 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(4):237-42. doi: 10.1136/flgastro-2013-100386
3. Stone H, Almilaji O, John C, et al. The Dedicated Iron Deficiency Anaemia Clinic: a 15-Year Experience. Frontline Gastroenterol 2020. doi: 10.1136/flgastro-2020-101470
4. 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 Gastroenterology 2020;7:e000403. doi: 10.1136/bmjgast-2020-000403
5. Wijayasekara C, Surgenor S, Hatcher C, et al. Clinical Risk Factors for Underlying Gastrointestinal Malignancy in Iron Deficiency Anaemia–Prospective Validation of The IDIOM Score. Cogent Medicine 2016;3 (1), 1213216. doi: 10.1080/2331205X.2016.1213216
6. National Institute for Health and Care Excellence (NICE). Suspected Cancer: Recognition and Referral 2015, updated Sep 2020. Last accessed 8 Nov 2020. Available from: https://www.nice.org.uk/guidance/ng12
7. Goddard A, James M, McIntyre A, et al. Guidelines for The Management of Iron Deficiency Anaemia. Gut 2011;60:1309-16. doi: 10.1136/gut.2010.228874.
8. Almilaji, O. Modelling the episodes of care for IDA patients in a secondary-care centre using continuous-time multistate Markov chain. Saudi J Gastroenterol. 2021;27(6).
9. Almilaji O, Snook J, Thomas P. The IDIOM App (ver 1.0): A Web-based Application to Predict the Risk of Gastro-Intestinal Cancer in Iron Deficiency Anaemia. [Software Medical Device, class I, MDD]. BU

Innovations Limited (BUI), UK 2020. Last accessed 7 9 2021. Available from: https://www.predict-gi-risk-in-ida.com

10. Snell KIE, Archer L, Ensor J, et al. External validation of clinical prediction models: simulation-based sample size calculations were more reliable than rules-of-thumb. J Clin Epidemiol 2021 Feb 14(135):79-89. doi: 0.1016/j.jclinepi.2021.02.011.
11. Van Calster B, McLernon DJ, Maarten Van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. BMC Medicine 2019;17(230) doi: 10.1186/s12916-019-1466-7
12. Kerr K. F., Brown M. D., Zhu K., et al. Assessing the Clinical Impact of Risk Prediction Models with Decision Curves: Guidance for Correct Interpretation and Appropriate Use. Journal of Clinical Oncology 2016;34(21). doi: 10.1200/JCO.2015.65.5654.
13. Kerr K. F., Brown M. D., Marsh T. L., et al. Assessing the Clinical Impact of Risk Models for Opting Out of Treatment. Medical Decision Making (MDN) 2019;39(2): 86-90. doi: 10.1177/0272989X18819479.
14. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. Diagnostic and Prognostic Research 2019;3(1):18. doi: 10.1186/s41512-019-0064-7.
15. Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. BMJ 2015; 350 doi: 10.1136/bmj.g7594.
16. Rockey DC. Gastrointestinal tract evaluation in patients with iron deficiency anemia. Semin Gastrointest Dis 1999;10(2):53-64. PMID: 10361896
17. Liu K, Kaffes AJ. Iron deficiency anaemia: a review of diagnosis, investigation and management. European Journal of Gastroenterology & Hepatology 2012;24(2):109-16. doi: 10.1097/MEG.0b013e32834f3140
18. Zhu A, Kaneshiro M, Kaunitz JD. Evaluation and treatment of iron deficiency anemia: a gastroenterological perspective. Digestive

diseases and sciences 2010;55(3):548-59. doi: 10.1007/s10620-009-1108-6

19. Rai S, Hemingway D. Iron Deficiency Anaemia: Useful Diagnostic Tool For Right Sided Colon Cancers?. Colorectal Disease 2005;7(6):588-90. doi: 10.1111/j.1463-1318.2005.00880.x
20. Surgenor S, Kirkham S, Parry S, et al. The Development of A Nurse-Led Iron Deficiency Anaemia Service In A District General Hospital. Frontline Gastroenterol 2014;5(3):219-23. doi: 10.1136/flgastro-2013-100385.
21. Almilaji O, Parry SD, Docherty S, Snook J. Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia. Sci Rep. 2021 Jun 22;11(1):13055. doi: 10.1038/s41598-021-92623-z.

Supplementary information

Table S 5-1 Model coefficients before/after applying different regularisation methods

	Full model[1]	Heuristic[2]	Bootstrap[3]	Penalised[4]	Lasso[5]
Intercept	-1.84	-1.80	-1.85	-1.77	-1.84
Sex (male)	0.939	0.917	0.919	0.903	0.894
Age (years)	0.057	0.056	0.056	0.055	0.049
MCV (fl)	-0.033	-0.032	-0.032	-0.032	-0.030
Hb (g/l)	-0.026	-0.025	-0.025	-0.025	-0.025

[1:] Standard binary logistic model

[2:] Heuristic shrinkage

[3:] Bootstrap shrinkage (500 bootstraps)

[4:] Penalized maximum likelihood using

[5:] Least absolute shrinkage and selection operator (Lasso)

Sample size considerations

Being a retrospective analysis of secondary data meant we did not have any control over the size of the external validation datasets. The number of outcome events in the Oxford and Sheffield datasets were 86, 36 respectively. Following the simulation-based sample size calculations for external validation of clinical prediction models [1], the anticipated precisions of performance measures were estimated based on the available number of outcome events in the external validation datasets, and on them both combined.

Using a fixed base probability of 0.084 (200/2390 from Dorset dataset), the simulation-based sample size calculations predicted an average 95% CI width for the discrimination value of around 0.16 when the standard deviation of linear predictor SD(LP) is 1 for 36 outcome events (n=429) as can be seen from Figure S 5-1. The figure was 0.10 for 86 outcome events (n=1024), and 0.08 for 122 outcome events (n=1452). Accordingly, to achieve a 95% CI width for the C-statistic < 0.1, 86 outcome events should be sufficient to achieve this precise performance measure. This means that the Oxford dataset can be used separately, or it can be combined with the Sheffield dataset to achieve more precise discrimination. A precision of discrimination performance < 0.1 cannot be achieved by using Sheffield dataset alone.

The simulation also showed that for the calibration slope (Figure S 5-2), a 95% CI width of 0.8 would be predicted at SD(LP) =1 for 36 outcome events, 0.6 for 86 outcome events, and 0.4 for 122 outcome events. Achieving a 95% CI width for the calibration slope < 0.2, would need about more than 500 outcome events. Thus, to achieve the most possible precise 95% CI width for the calibration slope, the two validation datasets must be combined.

Reference:

1. Snell, K., Archer, L., Ensor, J., et al. External validation of clinical prediction models: simulation-based sample size calculations were more reliable than rules-of-thumb. J Clin Epidemiol 2021 Feb 14(135):79-89. doi: 0.1016/j.jclinepi.2021.02.011.

Figure S 5-1 Average 95% confidence interval width for the C-statistic (discrimination) at different effective sample sizes comparing by SD(*LP*) at fixed base probability = 0.084

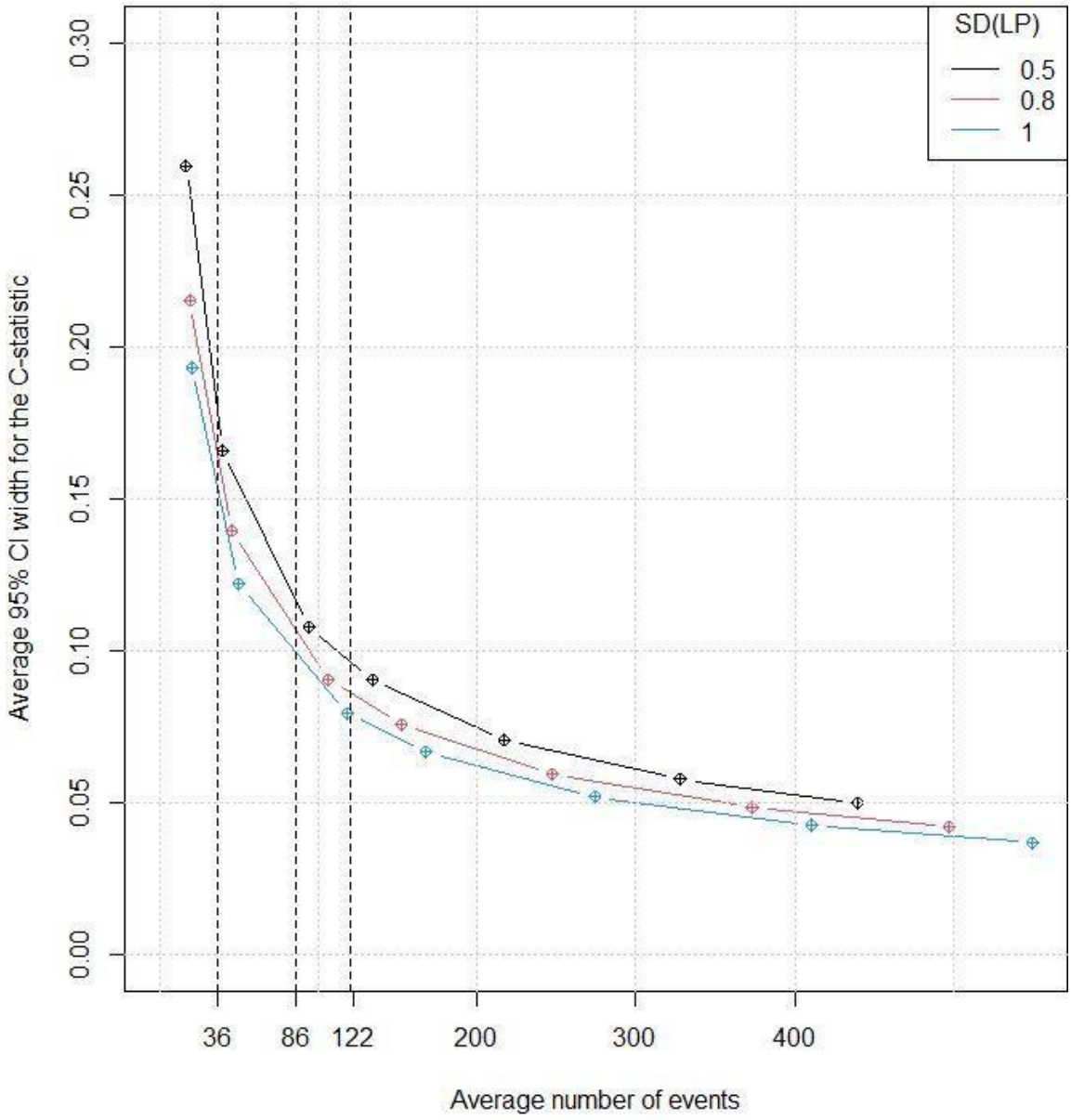


Figure S 5-2 Average 95% confidence interval width for the calibration slope at different effective sample sizes (based on number of events) comparing by SD(*LP*) at fixed base probability =0.084

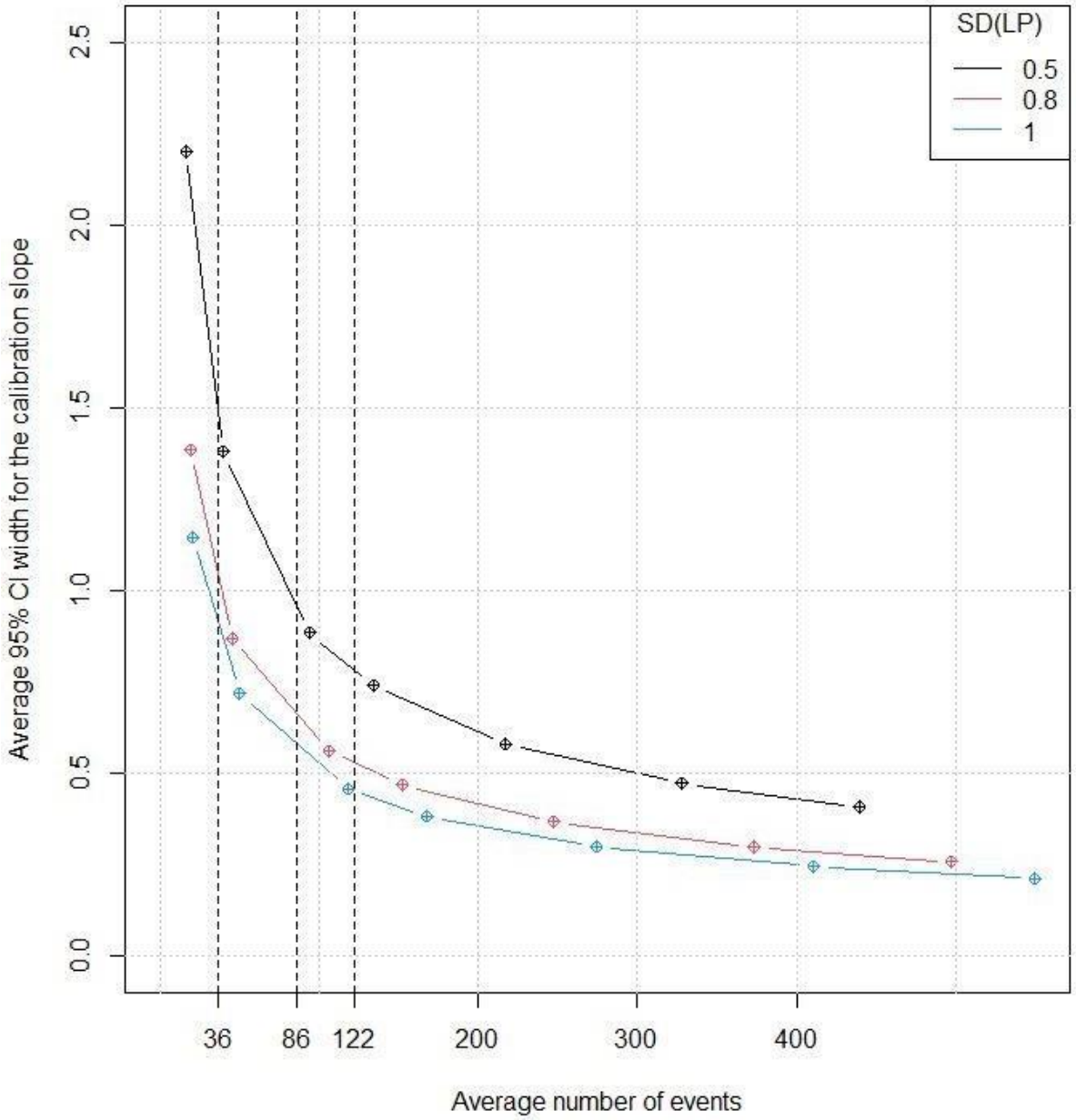


Table S 5-2 Normal ranges for Hb, MCV, ferritin, T.sat in each lab

	Dorset dataset	**Oxford dataset**	**Sheffield dataset**
Ferritin (ug/l)	Female: 13-150 Male: 30-400	Female: 10-200 Male: 20-300	31-400
T.sat (%)	15-45	16-50	Female: 15-45 Male: 15-50
Hb (g/l)	Female: 115-150 Male: 130-170	Female: 120-150 Male: 130-170	Female: 110-147 Male: 131-166
MCV (fl)	78-99	83-105	Female: 80.0-98.1 Male: 81.8-96.3

Figure S 5-3 Density plots show the distributions of continuous variables per GI presence/absence in each dataset using IDIOM model

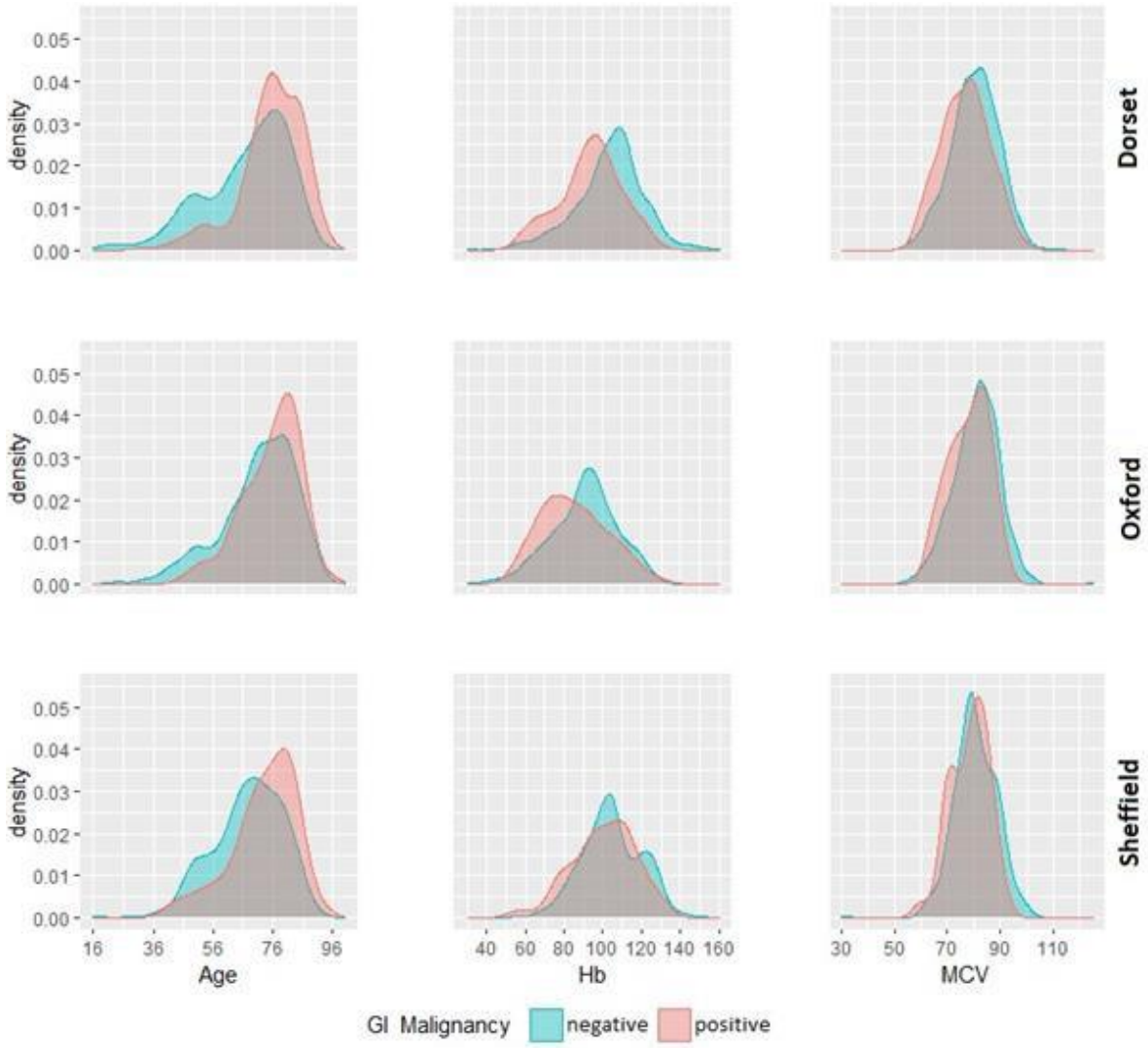


Figure S 5-4 Density plots show the distribution of estimated risks per GI presence/absence in each dataset using IDIOM model

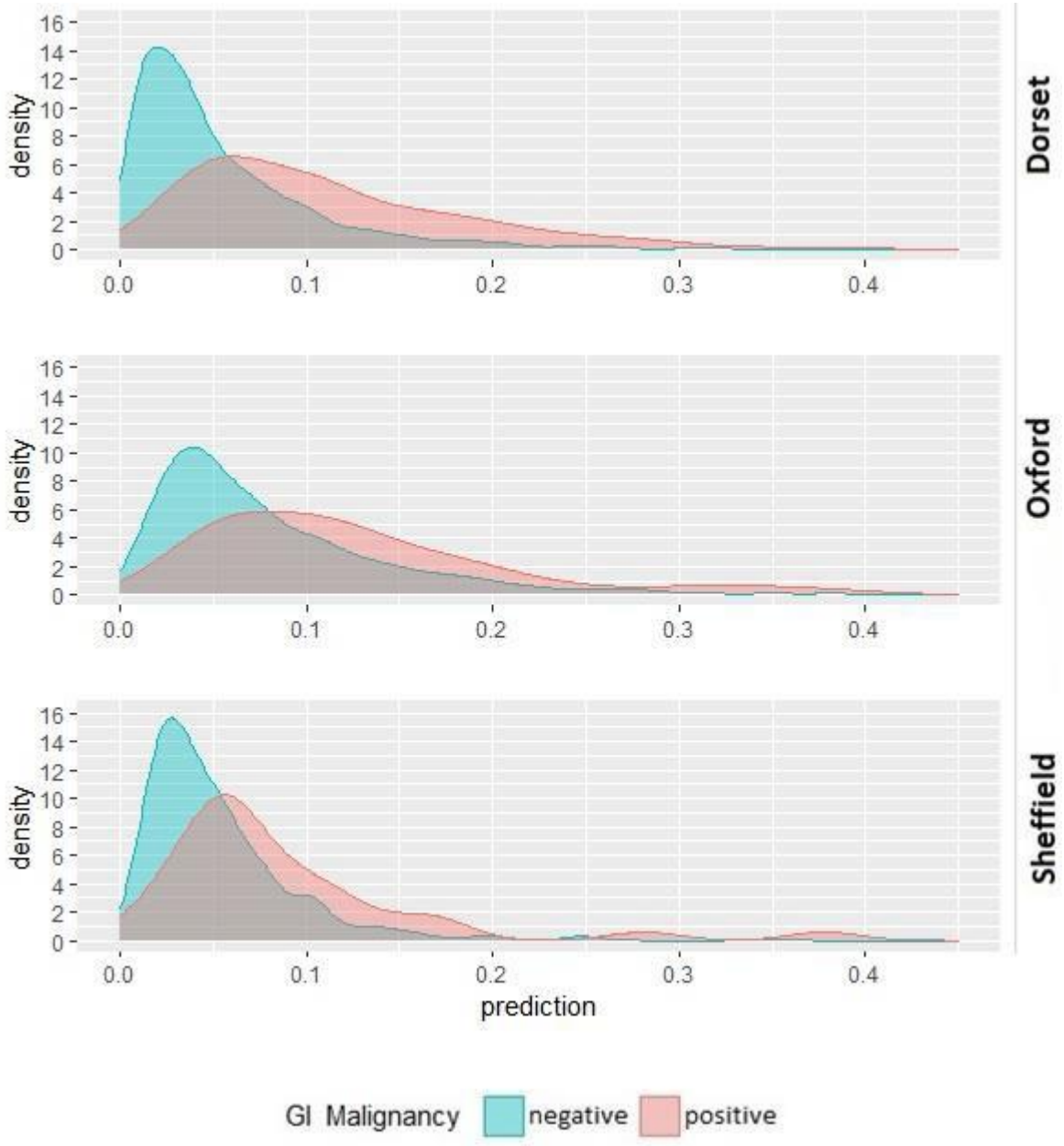


Figure S 5-5 Risk groups calibration in the combined validation dataset shows the relation between the estimated risks (on the x-axis) and the observed risks (on the y-axis)

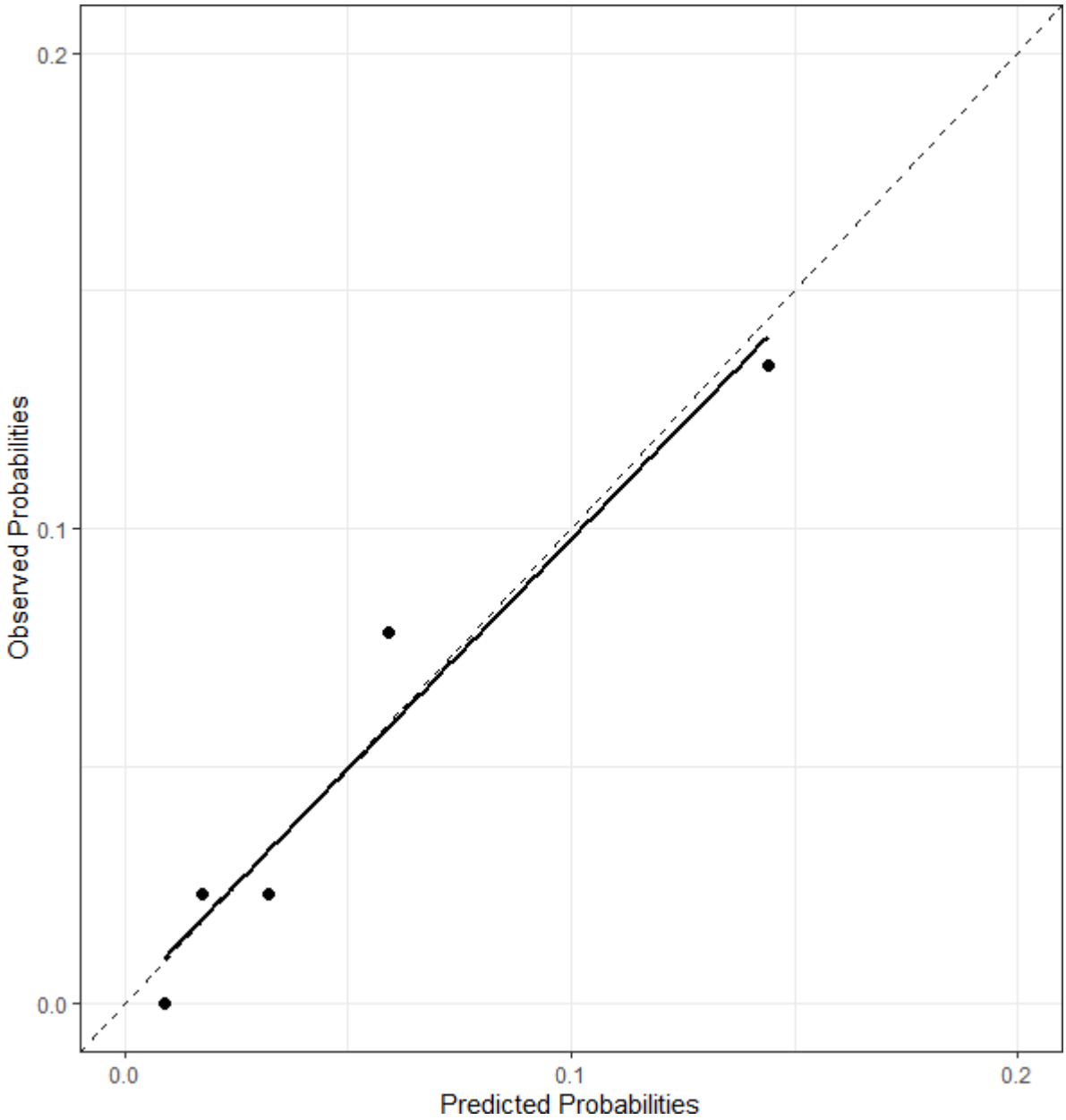


Figure S 5-6 Flexible calibration curve in Oxford dataset shows the relation between the estimated risks (on the x-axis) and the observed proportion of events (on the y-axis)

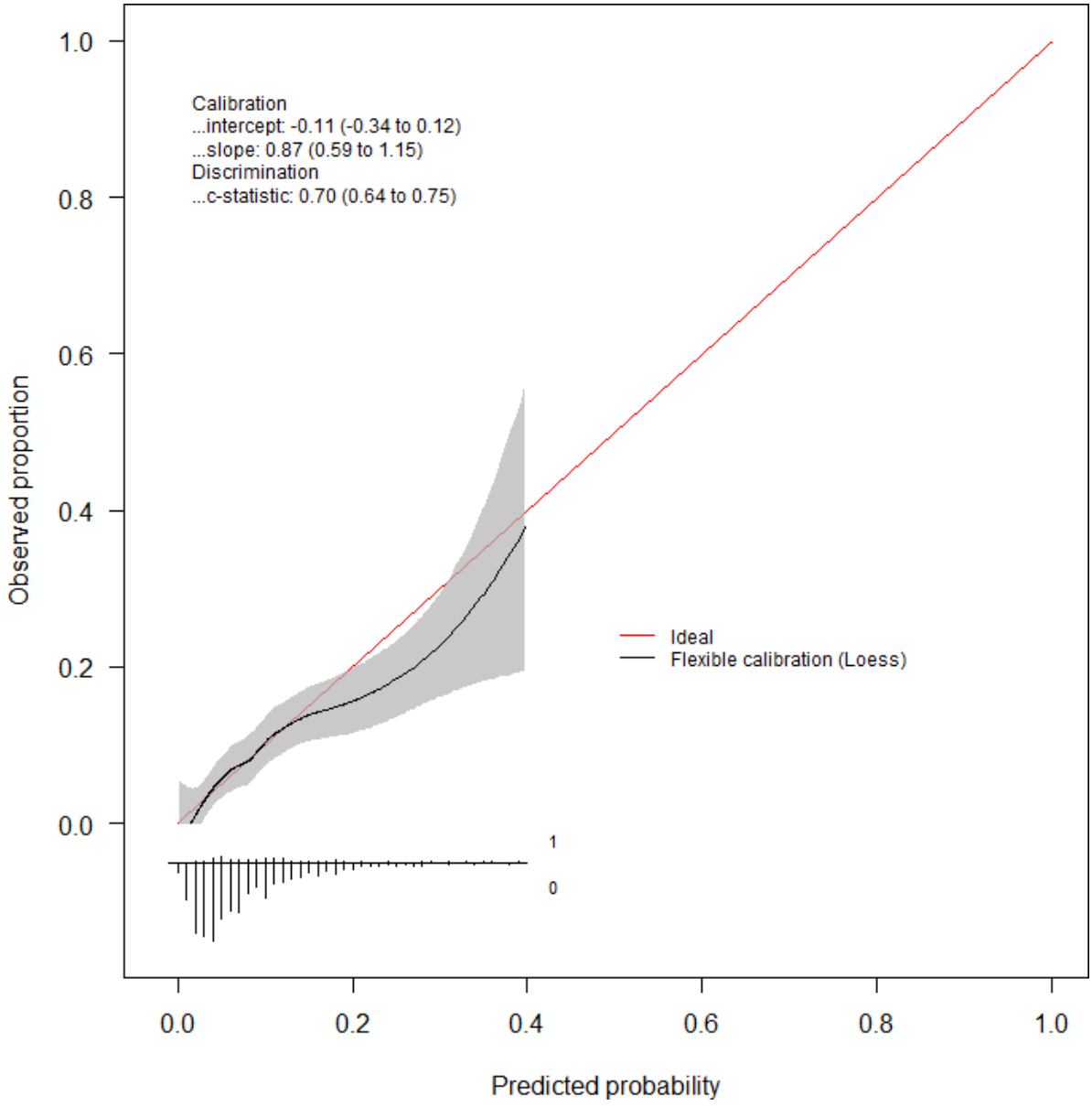


Figure S 5-7 Flexible calibration curve in Sheffield dataset shows the relation between the estimated risks (on the x-axis) and the observed proportion of events (on the y-axis)

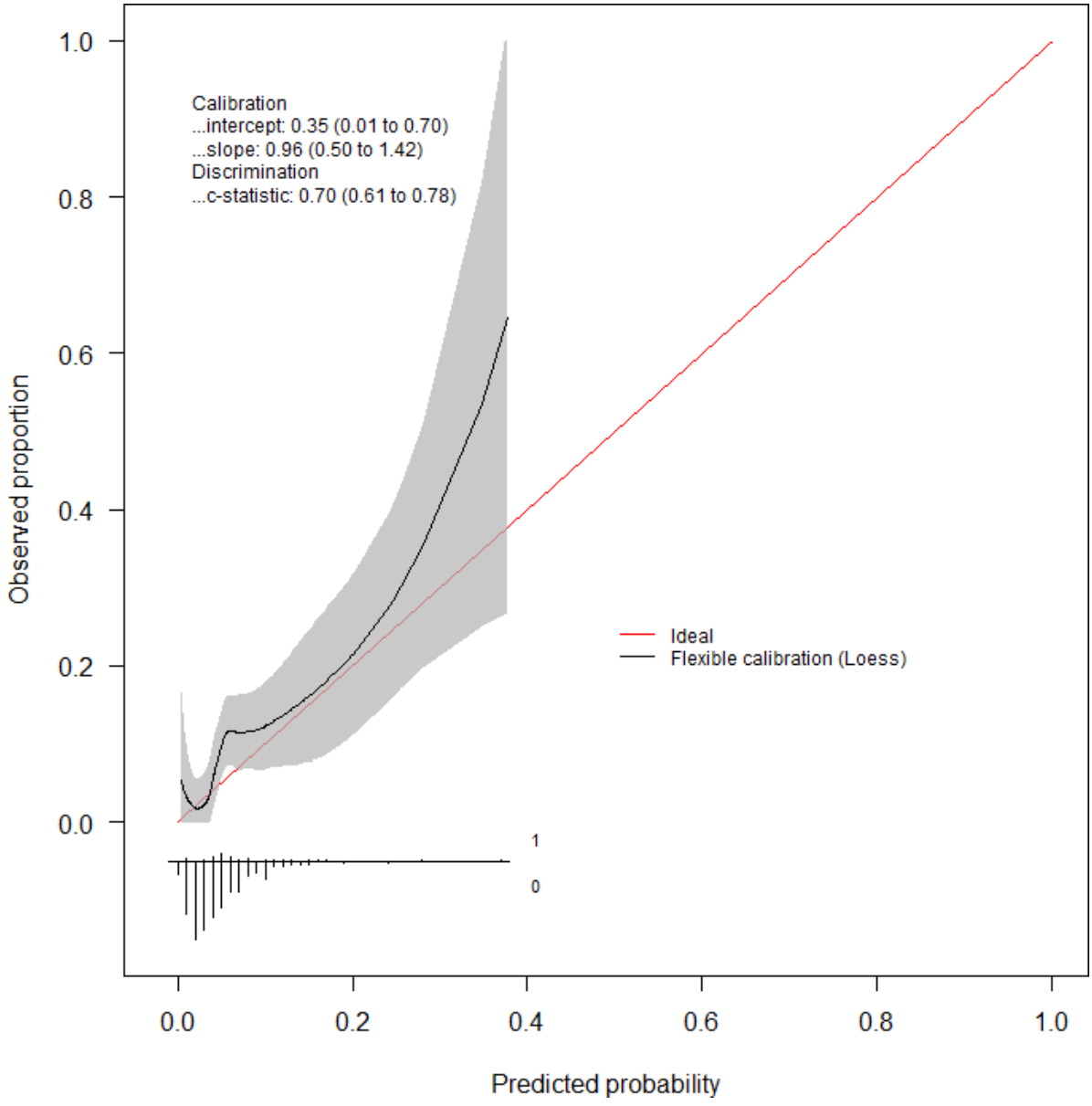


Figure S 5-8 Decision curve analysis for GI investigation using Oxford dataset. Grey line: penalised IDIOM model. Black line: "investigate no-one" strategy. Dashed line: "investigate all" strategy. The vertical axis displays standardized net benefit. The horizontal axis shows the risk thresholds

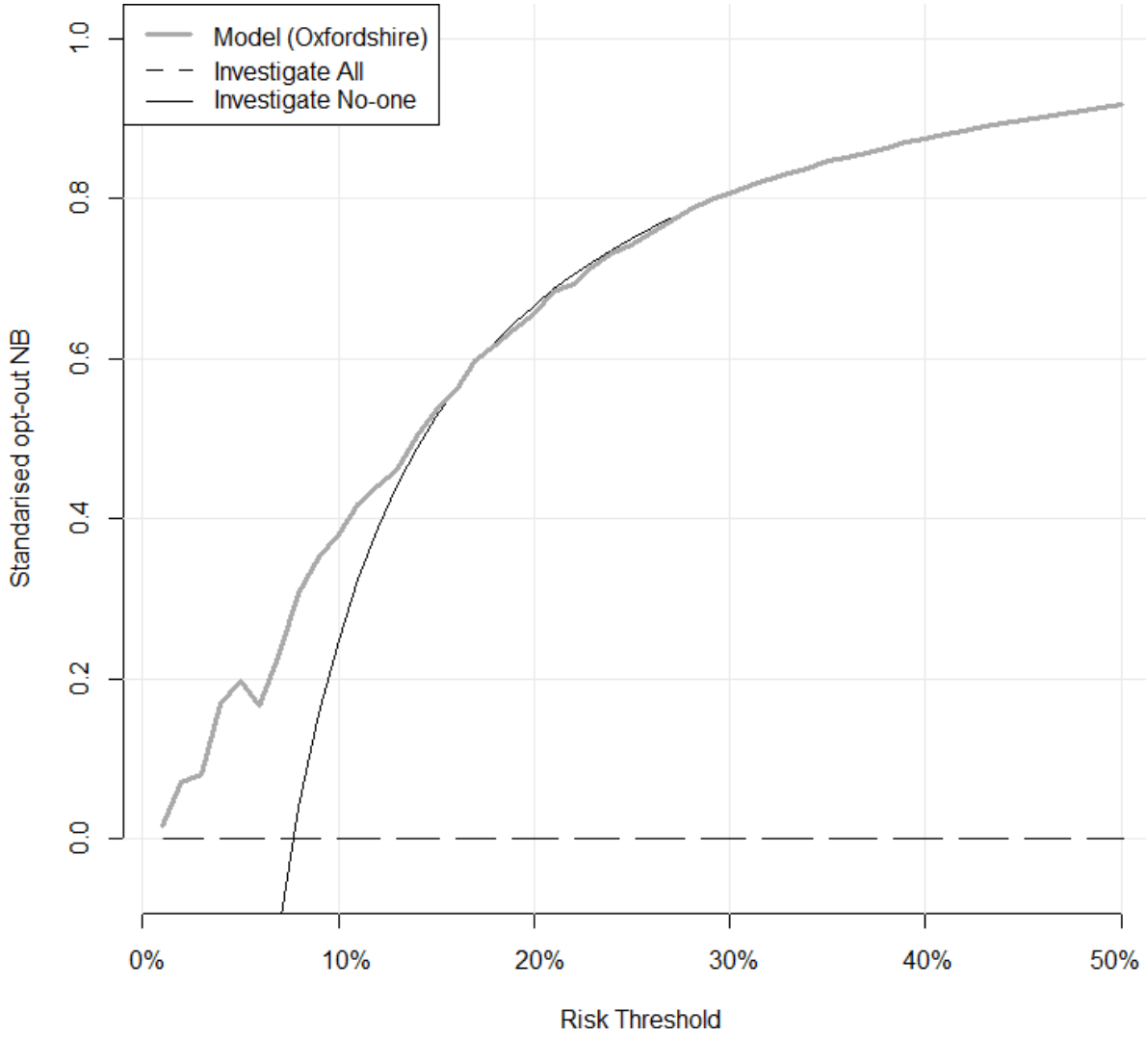


Figure S 5-9 Decision curve analysis for GI investigation using Sheffield dataset. Grey line: penalised IDIOM model. Black line: "investigate no-one" strategy. Dashed line: "investigate all" strategy. The vertical axis displays standardized net benefit. The horizontal axis shows the risk thresholds

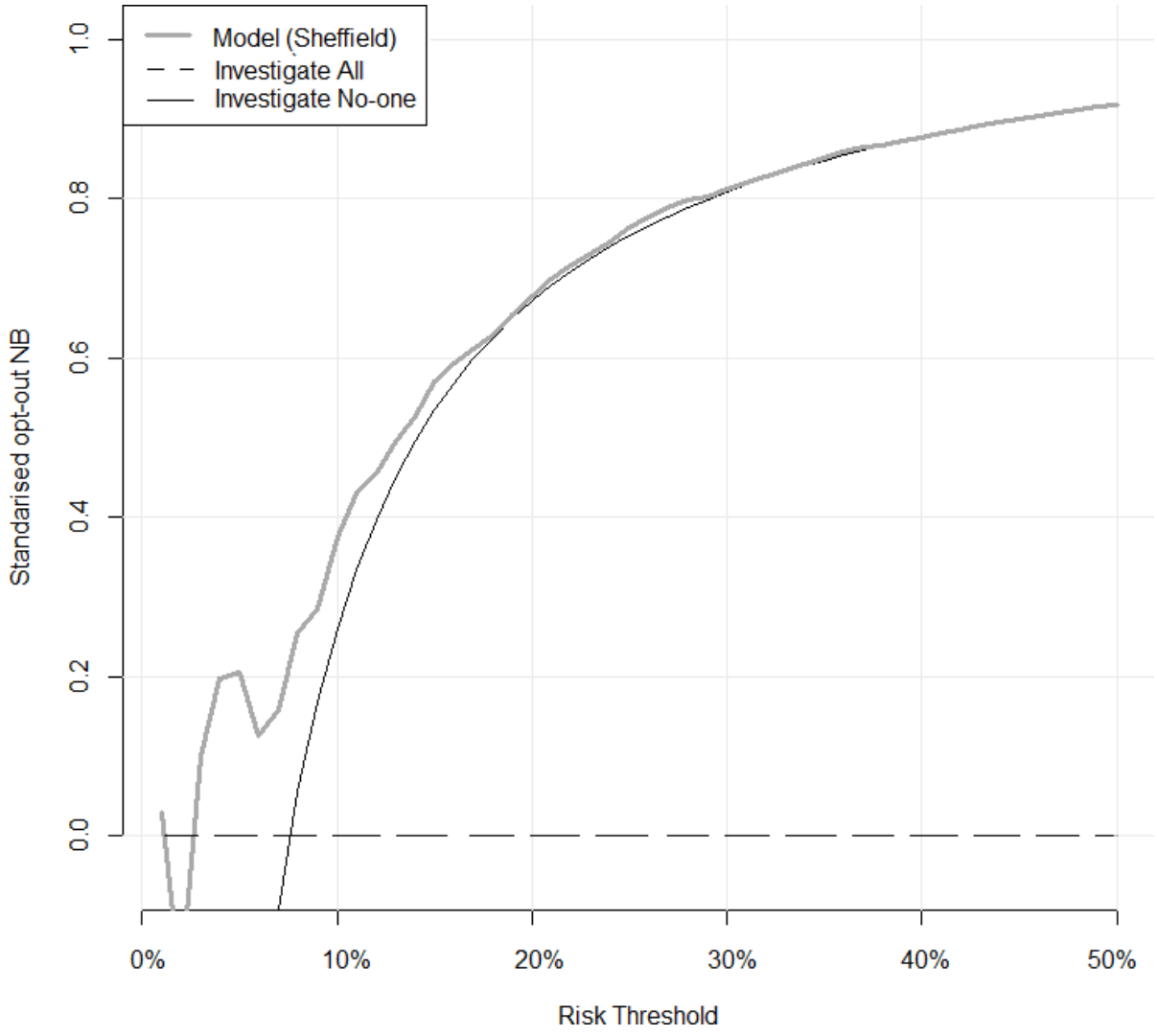


Figure S 5-10 Clinical impact curve for penalised IDIOM risk model using Dorset data, with 95% CIs constructed via bootstrapping. Of 1,000 patients, the heavy black solid line shows the total number who would be deemed low risk for each risk threshold. The blue dashed line shows how many of those would be true negatives. The vertical axis displays standardised net benefit. The two horizontal axes show the correspondence between risk threshold and cost:benefit ratio

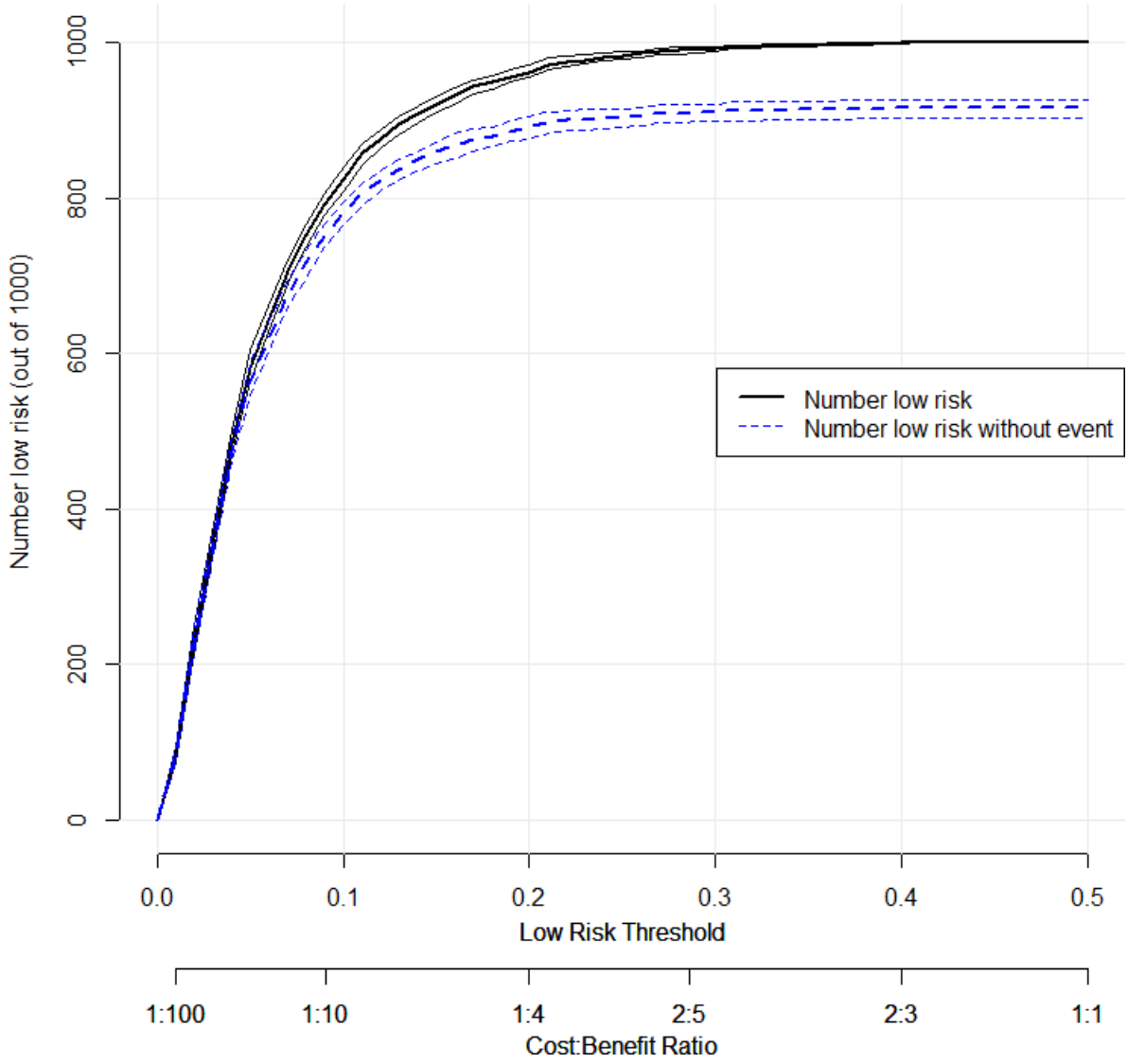


Figure S 5-11 Clinical impact curve for penalised IDIOM risk model using combined validation data from Oxford and Sheffield, with 95% CIs constructed via bootstrapping. Of 1,000 patients, the heavy black solid line shows the total number who would be deemed low risk for each risk threshold. The blue dashed line shows how many of those would be true negatives. The vertical axis displays standardised net benefit. The two horizontal axes show the correspondence between risk threshold and cost:benefit ratio

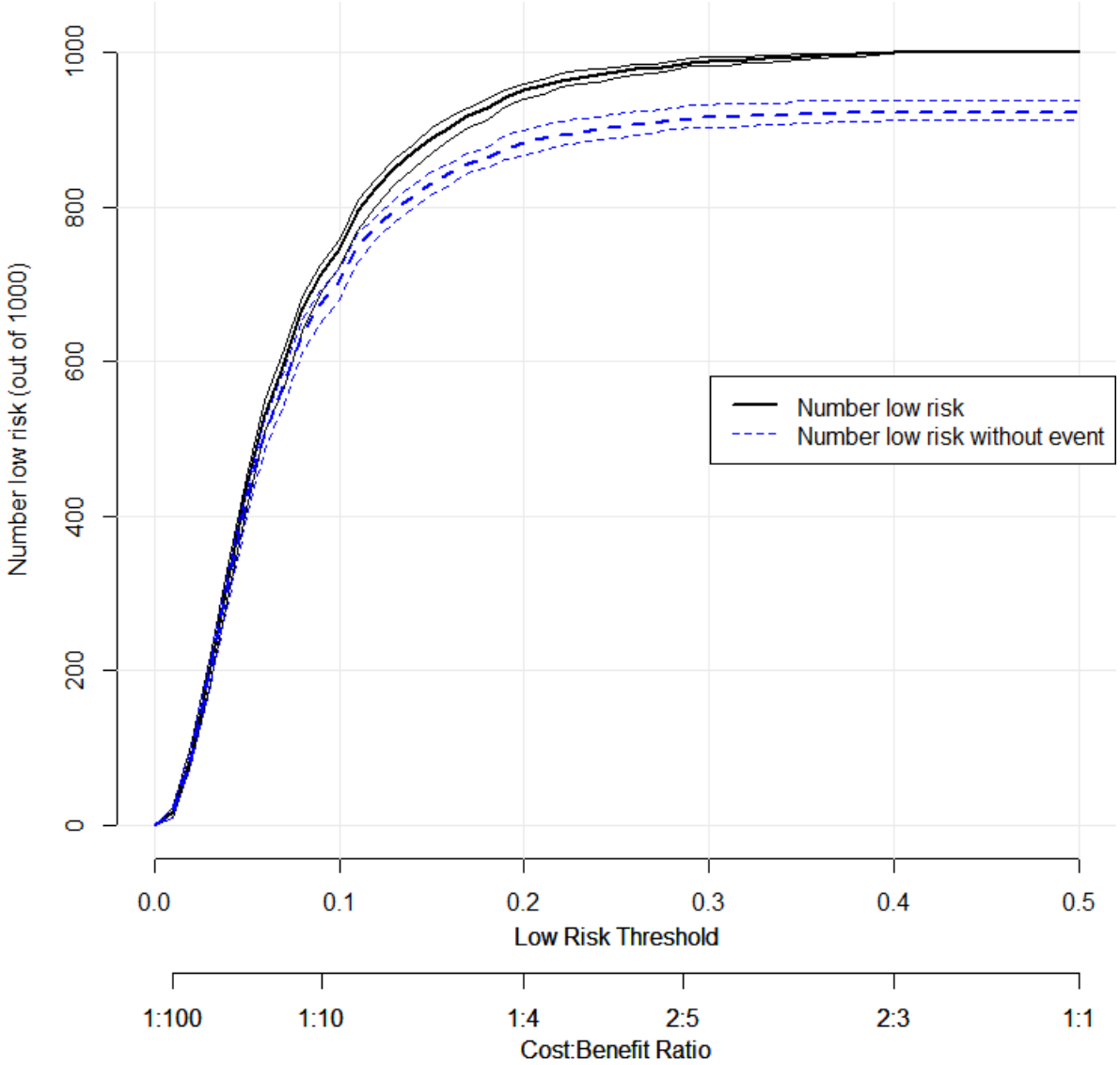


Figure S 5-12 Clinical impact curve for penalised IDIOM risk model using the Oxford dataset, with 95% CIs constructed via bootstrapping. Of 1,000 patients, the heavy black solid line shows the total number who would be deemed low risk for each risk threshold. The blue dashed line shows how many of those would be true negatives. The vertical axis displays standardised net benefit. The two horizontal axes show the correspondence between risk threshold and cost:benefit ratio

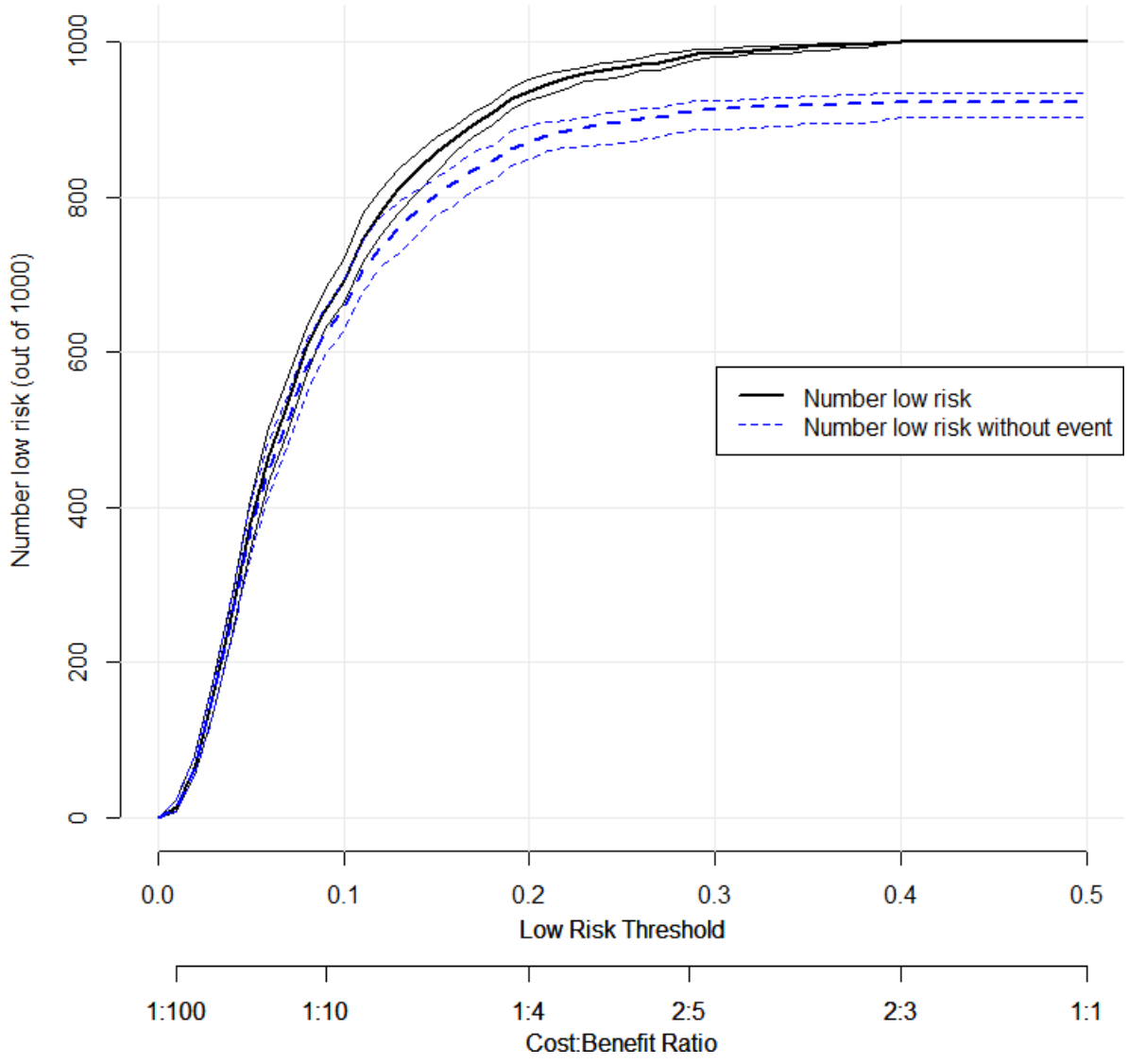


Figure S 5-13 Clinical impact curve for penalised IDIOM risk model using the Sheffield dataset, with 95% CIs constructed via bootstrapping. Of 1,000 patients, the heavy black solid line shows the total number who would be deemed low risk for each risk threshold. The blue dashed line shows how many of those would be true negatives. The vertical axis displays standardised net benefit. The two horizontal axes show the correspondence between risk threshold and cost:benefit ratio

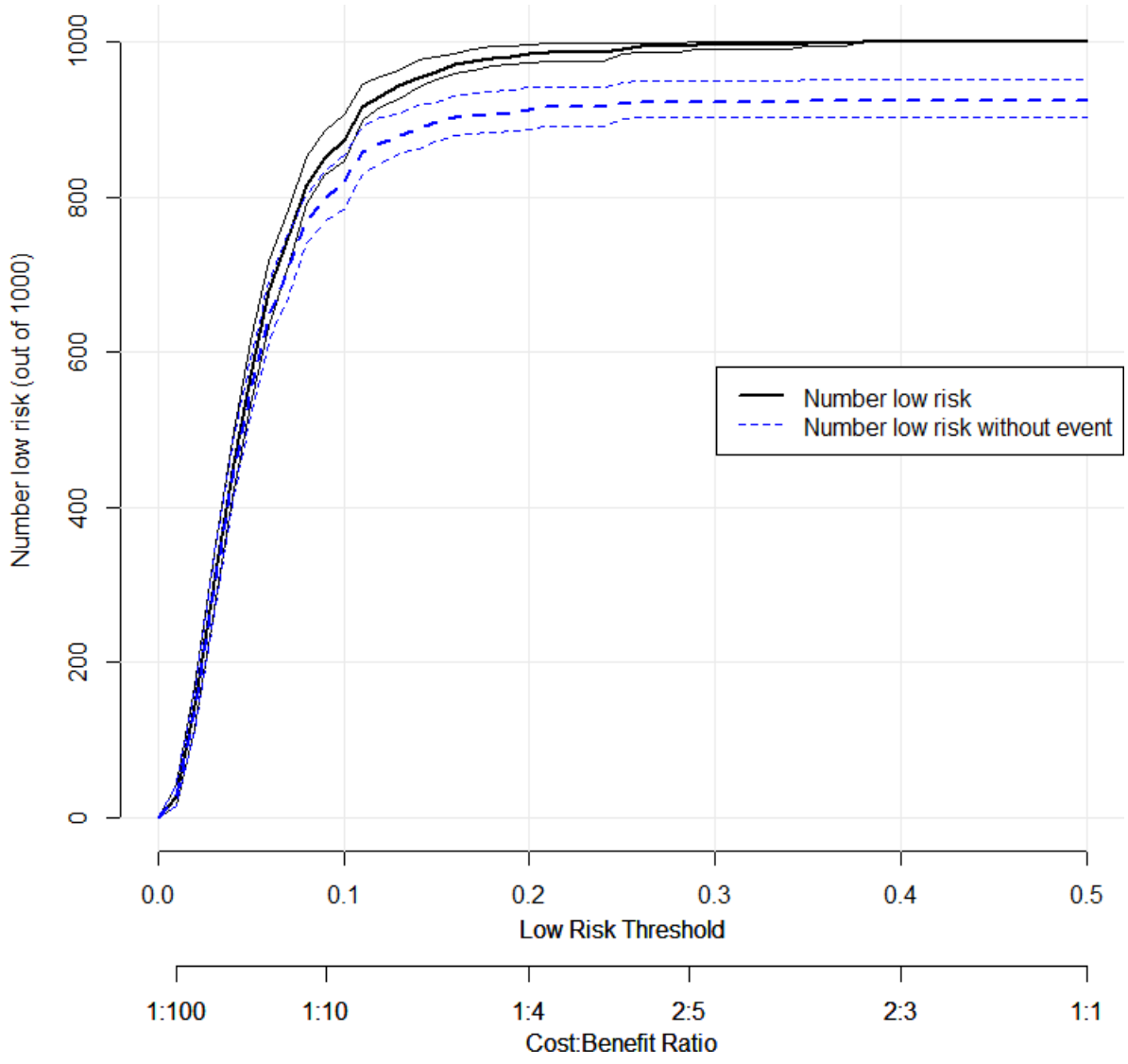


Table S 5-3 Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement

https://www.equator-network.org/reporting-guidelines/tripod-statement/

Section/ Topic	Item	Checklist Item	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	83
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	84
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	85
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	86
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	89
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	89
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	89
	5b	Describe eligibility criteria for participants.	89
	5c	Give details of treatments received, if relevant.	NA

Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	89
	6b	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	89
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	Explain how the study size was arrived at.	90
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	90
Statistical analysis methods	10c	For validation, describe how the predictions were calculated.	90
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	90-93
	10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	Provide details on how risk groups were created, if done.	90
Development vs. validation	12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	90
Results			
Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	93
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	94

	13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	94
Model performance	16	Report performance measures (with CIs) for the prediction model.	94-102
Model-updating	17	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	102-103
Interpretation	19a	For validation, discuss the results with reference to performance in the development data, and any other validation data.	102-103
	19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	104
Implications	20	Discuss the potential clinical use of the model and implications for future research.	104
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Attached file
Funding	22	Give the source of funding and the role of the funders for the present study.	14

Chapter 6 : Modelling the episodes of care for IDA patients in a secondary-care center using continuous-time multistate Markov chain

Open access original research article published at: https://www.saudijgastro.com/article.asp?issn=1319-3767;year=2022;volume=28;issue=2;spage=115;epage=121;aulast=Almilaji

Citation: Almilaji, O. Modelling the episodes of care for iron deficiency anemia patients in a secondary-care center using continuous-time multistate Markov chain. 2022. *Saudi J Gastroenterol*; 28:115-21. doi: 10.4103/sjg.sjg_387_21

Abstract

Background

Despite the high prevalence of gastrointestinal (GI) cancer in iron deficiency anemia (IDA), some IDA patients do not complete all the necessary GI investigations at the initial referral. As a result, existing cancers are diagnosed at a later referral with worse prognosis. The potential to detect GI cancer early depends on minimizing the delay time spent between the two consecutive referrals, where a patient did not complete investigations at the first referral, but at the second is diagnosed with positive GI cancer. This retrospective longitudinal study aims to highlight the proper methods to model these referrals.

Methods

Using anonymized data of 168 episodes of care for IDA patients at an IDA clinic in secondary care setting, continuous-time multi-state Markov chain is employed to determine the transition rates among three observed states for IDA patients at the IDA clinic, “incomplete investigations,” “negative GI cancer,” and “positive GI cancer” and to estimate the delay time.

Results

Once in the state of incomplete investigations, an estimated mean delay time of 3.1 years (95% CI: 1.2, 5) is spent before being diagnosed with positive GI cancer. The probability that the “positive GI diagnosis” is next after the state of “incomplete investigation” is 17% compared with 11% when it is followed in the state of negative GI cancer. Defining the survival as the event of not being in the state of “positive GI cancer,” the survival rate of IDA patients with negative GI cancer is always higher than those with incomplete investigations. Finally, being diagnosed with positive GI cancer is always preceded by the prediction of being considered “very high risk” at the earlier visit.

Conclusion

A baseline model was proposed to represent episodes of care for IDA patients at a secondary care center. Preliminary results highlight the importance of completing the GI investigations especially in IDA patients who are at high risk of GI cancer and fit enough to do the investigations.

Keywords: Endoscopy, episodes of care, gastrointestinal cancer, iron deficiency anemia, secondary care.

Introduction

The early detection of gastrointestinal (GI) cancer could lead to improve its prognosis. However, newly developed malignant tumors and some types of advanced cancers (right-side colorectal cancer) are asymptomatic or difficult to be picked up by the usual population screening programme (sigmoidoscopy, Fit test).[1] Accordingly, opt-in clinical investigations that target the 'at high-risk' population are necessary to detect these new or silent-type cancers. Due to the strong association between iron deficiency anaemia (IDA) and GI cancer,[2-6] and with the aim of managing IDA and investigating whether GI cancer is the underlying cause of any confirmed iron deficiency, a dedicated IDA nurse-led clinic was established under the supervision of the Gastroenterology Department at Poole General Hospital, UK in 2004.[4,7] The diagnosis of GI cancer is established by standard clinical investigations including gastroscopy, colonoscopy, CT scanning, and biopsy.[4]

Despite the high prevalence of GI cancer in IDA (8-10%),[8] and being a major trigger for urgent GI investigations,[9-14] due to informed patient preference, concurrent illness, or major co-morbidity including frailty, some IDA patients do not complete all the necessary GI investigations.[4] Consequently, cancers that already existed during the time at which patients did not complete their investigations are diagnosed at later referrals with worse prognosis. Many factors may influence any patient's re-referral to the clinic such as new

symptoms including rectal bleeding, weight loss, stomach pain, and so on and also, being a recurrent IDA patient who is willing and fit enough to undergo the GI investigations.

The time spent by a confirmed IDA patient between two consecutive referrals to the clinic, where, at the first episode of care the required GI investigations were not completed, and at the second the patient is diagnosed with positive GI cancer is referred to as the 'delay time'. The potential to detect GI cancer early depends on minimizing this delay time. To predict the risk of GI cancer in patients with confirmed IDA, a binary multivariable logistic model was previously built and internally/externally validated based on four simple variables: age, sex, hemoglobin concentration (Hb), and mean cell volume (MCV)– the IDIOM model (Iron Deficiency as an Indicator of Malignancy).[4,15] Based on the predicted cancer risks that were derived from this model, IDA patients were stratified into five risk groups in which the lowest risk group (ultra-low risk) represents the lower half of the first quarter of positive predictive values with negative predictive values =100%, and the highest risk group (very-high risk) represents the fourth quarter of positive predictive values.

Due to the small size of the available multi-state data, in which only 168 episodes of care were found in the admission history at the IDA clinic for 83 patients with only four positive GI cancer cases at the subsequent episodes of care, the leading focus of this study is on gaining insights into the proper methods of modelling the episodes of care for IDA patients at the IDA clinic, and not on making inference from the preliminary results of applying these proposed methods on such small sample size. Therefore, when enough data becomes available in the future from a subsequent temporal period at the same clinic and/or from other similar secondary-care centers, a large-scale study can make use of the suggested methodology in this study to estimate the delay time, and to examine whether being stratified in ultra-low risk or very-high risk group by the IDIOM score at the earlier episode of care could

lead to being diagnosed with positive GI cancer at the following episode of care.

Methods

Study population

A total of 2788 patients with no other neoplasm, and with confirmed iron deficiency were referred to Poole hospital IDA clinic during the period of 2004-2018. Confirmed iron deficiency was defined by transferrin saturation <15% and/or serum ferritin less than the lower laboratory limit of normal at the time of the analysis. The anonymized secondary data for each referral, per patient, included:

- Patient ID
- Sex
- Age
- Blood hemoglobin concentration (Hb)
- Mean cell volume (MCV)
- Iron studies (transferrin saturation and serum ferritin)
- Date of the visit(s) to the IDA clinic
- The outcome of the GI investigation (positive/negative GI cancer)
- Indicator of the GI investigations' completion

GI Investigations were considered "complete" if the upper GI tract had been examined by gastroscopy, and the colon had been fully imaged either by colonoscopy or CT colonography.[4]

Statistical analysis

Usually patients are seen at intermittent referral visits in the IDA clinic, at which admission information is collected, but information from the periods

between visits is not available. The admission history (or the outcomes of episodes of care), for any IDA patient, comprises being observed either in the state of incomplete investigations, in the state of positive GI cancer, or in the state of negative GI cancer but never in any more than one state at one time; these states are finite disjoints states. Because the durations between the consecutive admissions to the clinic are irregularly spaced, a continuous-time multi-state Markov chain was appropriate to model these states, to determine the transitions rates between states, and to estimate the delay time.

Due to their ability to represent repetitive events, and time, Markov chains have been used intensively to model transition rates in clinical settings. In particular, Markov chains are frequently used to model disease progression.[16] Markov models are often developed to represent random processes that evolve over time.[17] These random processes satisfy the Markov property of “memorylessness”.[18] That is, the state of the process at a future time, given the previous history of the process up to the present time, depends only on the present-time state. These models assume that an entity is always in one of a finite number of discrete states, called Markov states, and all events are represented as transitions.[19] IDIOM score was used to predict the GI cancer risk for each patient and to stratify the patient per visit in the different risk groups based on the threshold proposed in Almilaji et al.[15]

Specifying the baseline model

The patient clinic admission history was modelled in a three-states continuous-time Markov model [Figure 6-1], through which the IDA patient can be moved in. These observed states are: S1) incomplete investigations, S2) negative GI cancer, and S3) positive GI cancer. “Death” state was not included in the model due to the totally missing information about this event, and because the time spent in states S1 or S2 is independent of any transition after S3. The time of observation refers to the last time the patient is seen at the clinic per referral and is used as surrogate time for the diagnosis time.

Time interval between any pair of consecutive visits per patient is measured in years.

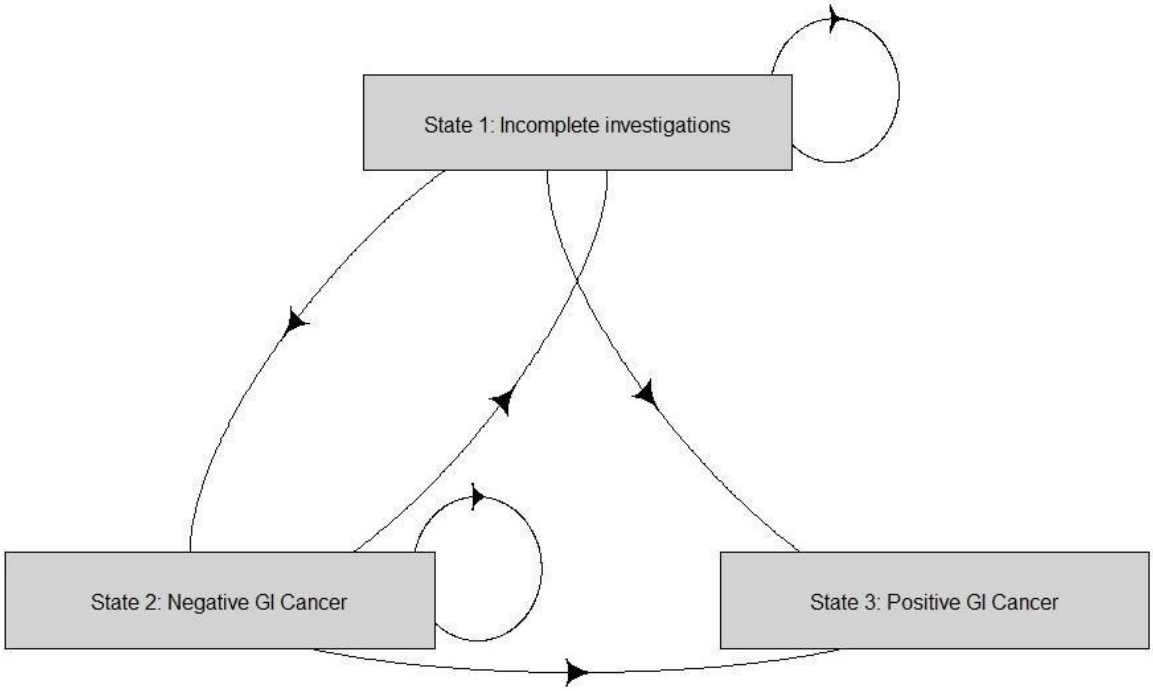


Figure 6-1 Markov-state diagram. The rectangles represent states, arrows represent transitions between states. Arrows leading from a state to itself indicate that the patient can remain in that state in consecutive cycles

Model assumptions

- For each instant of time t, for each pair of states the probability of an event at time t+1 depends exclusively on the actual state of the process and not on the previous states (Markov property).
- Transition probabilities only depend on the difference t between s and s + t and not on the actual times (s, s + t) that is the Markov model is homogeneous.
- As any clinical diagnosis is based on complete investigations, positive and negative GI cancer are assumed to be 100% accurate. So, no misclassification is proposed in this model.

- Positive GI cancer stage is an absorbing state as the patient cannot go back to the other states once it enters this absorbing stage. Once a patient is diagnosed with positive GI cancer, he/she will be transferred from the IDA clinic to another specialist clinic to start receiving the cancer treatment.
- Though some patients might totally avoid the GI investigations, in this analysis, “non-investigations” is regarded as a subset of incomplete investigations.
- The observation times vary either randomly and independently of the current outcome of the investigations, or according to primary care policies in which IDA patients with new signs of GI cancer are re-referred to the clinic. Hence, observation times are assumed to be non-informative sampling times.[20]

Intensity matrix

The tendency of a patient to make a transition from one state to another is described by the rate of transition (transition intensity). Transition rates ($\mathbf{q_{ij}}$) are elements of an intensity matrix Q, in which at time t > 0, it is given by:

$$Q = \begin{bmatrix} -(\mathbf{q_{12}} + \mathbf{q_{13}}) & \mathbf{q_{12}} & \mathbf{q_{13}} \\ \mathbf{q_{21}} & -(\mathbf{q_{21}} + \mathbf{q_{23}}) & \mathbf{q_{23}} \\ 0 & 0 & 0 \end{bmatrix}$$

The proposed model is governed by this transition intensity matrix. The transition rate represents the number of occurrences of an event for a given number of patients per unit of time and is similar to an instantaneous velocity. It can take any value in the range [0, ∞]. The rows sum, in this matrix, to 0. The diagonal entries are defined as minus the sum of all the other entries in the row. It is important to remember that the data are assumed to represent snapshots of the process at arbitrary times and fitting the model is a process of finding values of the four unknown transition intensities: q12, q13, q21, and

q23, which maximize the likelihood. Transition probabilities for any time t, calculated by taking the matrix exponential of the scaled transition intensity matrix:

$$P(t) = e^{tQ}$$

The final row is all zeroes in this Q matrix because positive GI cancer is an absorbing state and there are no transitions back to the other states. Inevitably, when insufficient data is used, the parameters of the proposed model (transition intensities) cannot be identified. Hence, given the small size of the data, the proposed model in this study was built as a simple model with no covariates.

As this study was retrospective analysis of anonymized secondary data, no patient was involved. R (version 3.6.1), RStudio (version 1.2.5001), and msm package were used to run the statistical analyses and to produce the graphs. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines[21] were used to ensure the reporting of this study.

Results

Patients

Patients started to be re-referred the clinic in 2008 at which the number of returning visits has started to increase gradually. The median time between any two consecutive referrals for all patients at the clinic was about 3 years. The median age of the 83 patients' cohort was 70 years (IQR: 60–77). Despite the four positive GI cancer cases at the subsequent episodes of care were all for male IDA patients, female patients were more likely to re-visit the clinic than male patients (Female/Male sex ratio: 2.5 (= 59/24)) as can be seen from Figure 6-2.

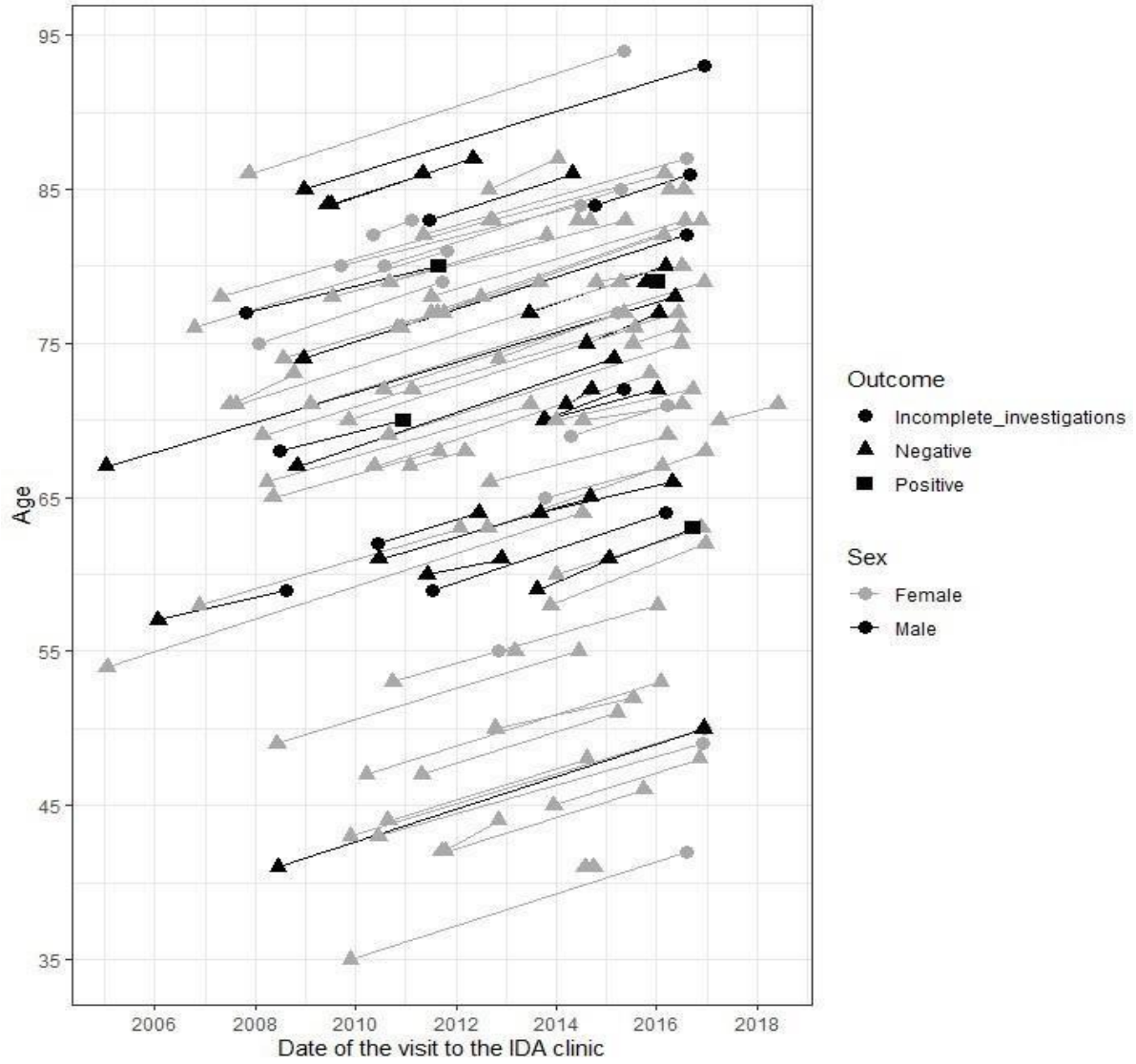


Figure 6-2 Patients' admission history to the IDA clinic during the study period 2004–2018

During the study period, there were 2873 episodes of care. About 2788 of these represent the first episodes of care for every patient, in which, 393 patients had incomplete investigations, 2194 diagnosed with negative GI cancer, and 201 diagnosed with positive GI cancer. Of the patients who had negative GI cancer or incomplete investigations, 83 had been re-referred to the clinic for the second time. About 18 of these patients did not complete investigations, 62 were negative GI cancer, and three were positive GI cancer.

Two of these 83 patients whose previous diagnoses were negative have been re-referred to the IDA clinic for the third time in which one was diagnosed with positive GI cancer and one with negative GI cancer as can be seen from the following patients' flow chart [Figure 6-3].

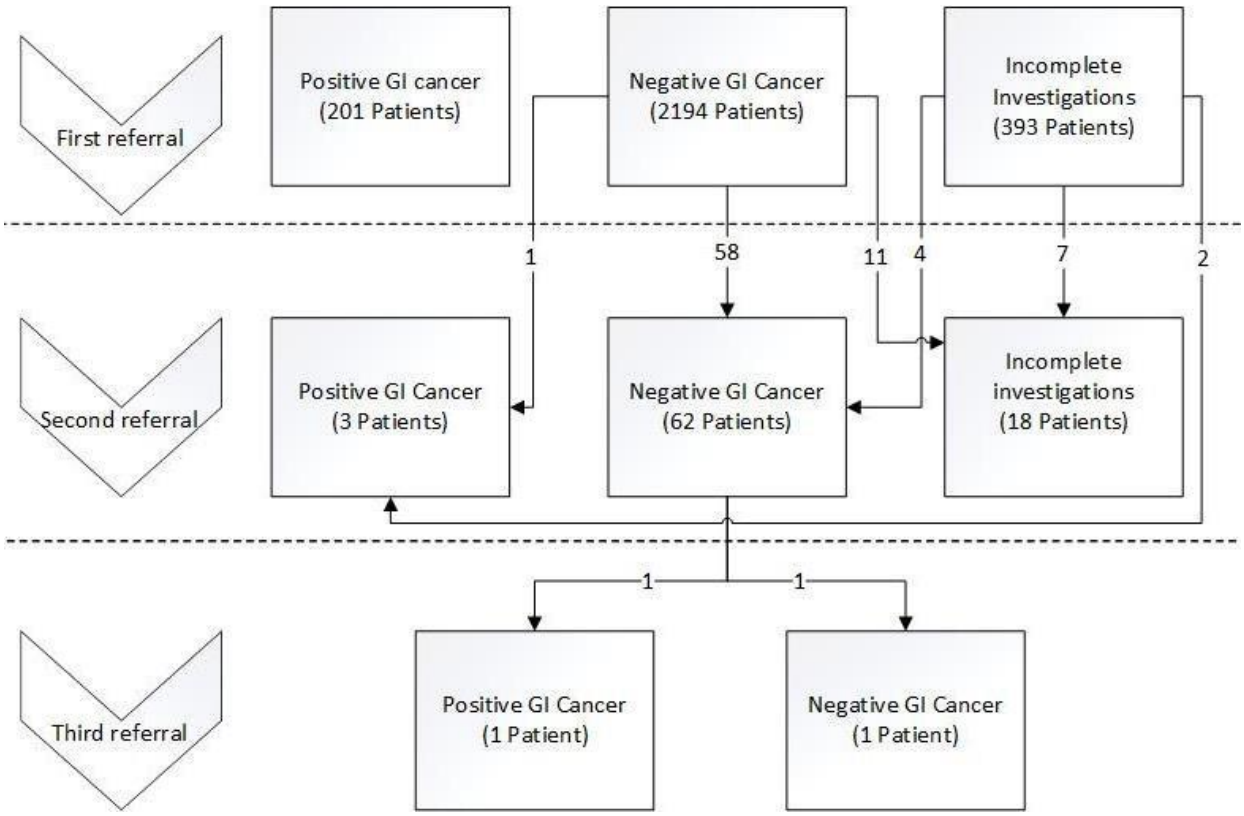


Figure 6-3 Flow chart of patient' states at the IDA clinic during the study period 2004–2018

To summarize the multi-state data in this study, a frequency table of pairs of consecutive states that counts for all patients, the number of times a patient had an observation of one state followed by an observation of another state is presented [Table 6-1]. Thus, out of the four GI positive cases, two came from state 1 (incomplete investigations), and two from state 2 (negative GI cancer) as can be seen from Table 6-1.

Table 6-1 Frequency table of consecutive states pairs

		To		
		Incomplete investigations	**Negative GI cancer**	**Positive GI cancer**
From	Incomplete investigations	7	4	2
	Negative GI cancer	11	59	2

Transition intensities estimates and 95% confidence intervals (CI)

The maximum likelihood estimates of the unknown parameters and 95% confidence intervals (CI) is given in Table 6-2.

Table 6-2 Estimated transition Intensities

Transition Intensities	**Estimates (95% CI)**
State 1–State 1	-0.184 (-0.46, -0.07)
State 1–State 2	0.153 (0.05, 0.45)
State 1–State 3	0.031 (0.01, 0.14)
State 2–State 1	0.066 (0.03, 0.14)
State 2–State 2	-0.075 (-0.15, -0.04)
State 2–State 3	0.008 (0.002, 0.03)

From the estimated intensities of the fitted model in Table 6-2, it can be seen that the rate of moving from “incomplete investigations” to “positive GI diagnosis” (0.031) is higher than that of moving from “negative GI diagnosis” to “positive GI diagnosis” (0.008). Patients are five times (0.153/0.031) more likely to be negative GI cancer than positive GI cancer at a later visit to the clinic (transitions from state 1). After being diagnosed with negative GI cancer

moving into the state of being not investigated state is eight times (0.066/0.008) more likely than the progression into positive GI cancer.

Once in the state of negative GI diagnosis, an estimated mean of 13.4 years (95% CI: 6.8, 26.2) is spent in the state of negative GI diagnosis before being diagnosed with positive GI cancer or moved into the state of being with incomplete investigations. And the probability that the “positive GI cancer” is next after the state of “negative GI cancer” is 11%. Once in the state of incomplete investigations, an estimated mean of 5.4 years (95% CI: 2.2, 13.4) is spent in the state of incomplete investigations before being diagnosed with negative or positive GI cancer. The estimated mean delay time was 3.1 years (95% CI: 1.2, 5). And the probability that the “positive GI diagnosis” is next after the state of “incomplete investigation” is 17%.

Survival plot

Defining the survival as the event of not being in the state of “positive GI cancer,” the 10-year survival probability for IDA patients with negative GI diagnosis is approximately 0.87, as opposed to 0.79 with incomplete investigations. Accordingly, the survival of IDA patients with negative GI diagnosis is always higher than those with incomplete investigations as can be seen from Figure 6-4.

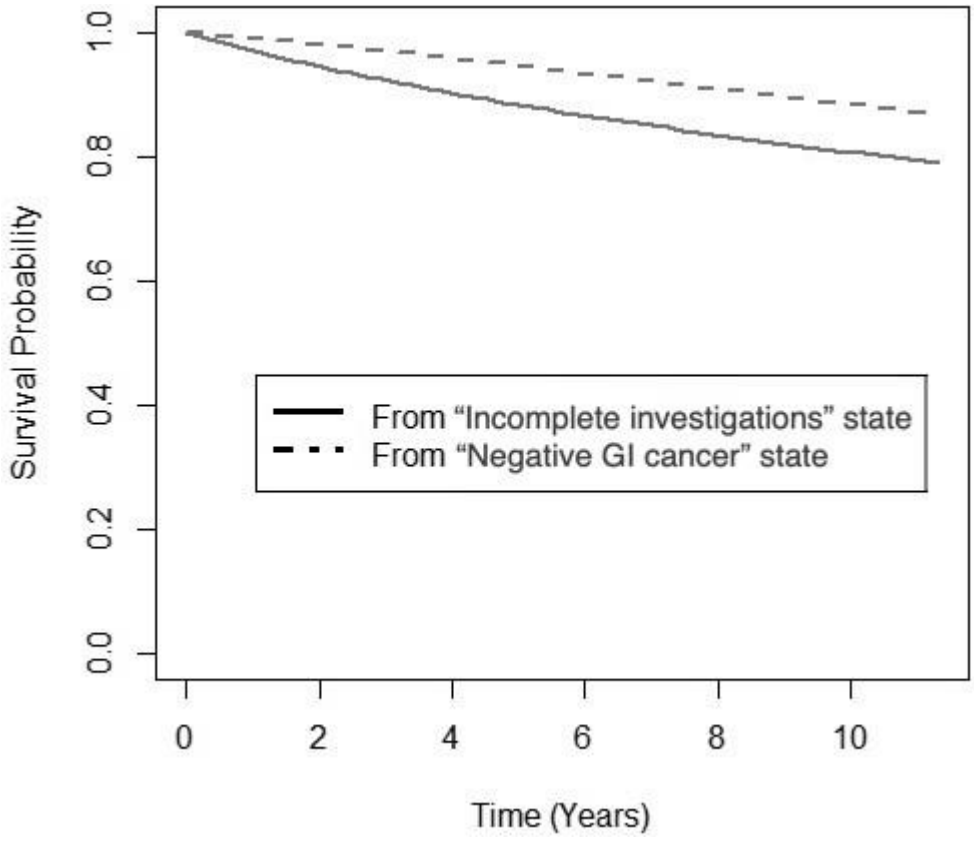


Figure 6-4 Survival plot. Survival is defined as not being in the state of “positive GI cancer”

IDIOM risk groups

At the following visits, for all patients who have completed their investigations, no difference was found between the observed GI cancer risk that was 6% (4/67) and the 8% predicted risk by IDIOM. A preliminary conclusion could be that that recurrent IDA is not a risk factor for GI cancer. Interestingly, the four patients who have been diagnosed with positive GI cancer were predicted by IDIOM score to be in the very-high risk group at the earlier visits. Also, all the patients who were predicted to be in the lowest risk group at the earliest visits and complete their investigations at the follow-up visits were diagnosed with negative GI cancer.

Discussion

About 14% of the patients who were referred to the IDA clinic did not complete their investigations at the first referral to the clinic compared with 79% diagnosed with negative GI cancer and 7% with positive GI cancer at the same first referral. About 21% did not complete their investigations at their subsequent referrals to the clinic compared with 74% diagnosed with negative GI cancer and 5% with positive GI cancer at the following referrals.

Applying the proposed methods on the available data showed that the transition rate of moving to positive GI cancer is higher when patients are observed in incomplete investigations state than negative GI cancer. The average delay time in “incomplete investigations” for IDA patients is about 3 years, and the probability that a positive GI cancer is followed by the state of incomplete investigations was 17% compared with 11% when it is followed by the state of negative GI cancer. Another finding was that the survival of IDA patients with incomplete investigations was always lower than those with negative GI cancer despite the fact that the waiting time in the state of “negative GI cancer” was about double the time of the delay time. Finally, being diagnosed with positive GI cancer always preceded by the prediction—according to IDIOM score—of being considered very high risk at the earlier visit. Nevertheless, as mentioned earlier, these former findings are preliminary results only and should always be reported within the context of the available small-size data and interpreted with caution especially that only two patients developed cancer from the group of incomplete investigations and the other two developed from previously negative diagnosis group. The small numbers of patients have resulted in wide confidence intervals for the estimates.

The limitations of this study include the inability to increase the size of the sample, and accordingly restricting the analysis to a baseline model of the transitions between consecutive admissions. However, for any future large-scale studies using the methodology proposed in this study, we should take

into consideration the following issues that became apparent while developing the model:

1. Transition rates might be dependent on patient-related variables such as sex, age, and other pathologies including inflammatory disease, celiac disease, adenoma, and so on. For any future model to be accurate, the effects of these covariates on the transition rates should be addressed by using a proportional intensities model.

2. In the developed baseline model, there was no differentiation between the events of “incomplete investigations” and “no investigations”. A question about whether being observed with partial or no investigations could affect the transition rates to the positive GI cancer state differently must be answered. If a variance is found, a separation between these two states should be adopted in any future model.

3. One of the assumptions in this study was that a negative GI cancer is always accurate because it is based on full clinical investigations, and thus there was no account for any misdiagnosis margin. A future comprehensive model must investigate and support this claim.

4. This study implicitly assumed that in those patients who were diagnosed with positive GI cancer at the subsequent referral to the clinic after not completing the investigations at earlier referral, the GI cancer had already existed at the time of the first referral. However, high-grade aggressive GI cancer could have an onset time between the consecutive referrals. One way to compensate for this fact is to include the GI cancer grade and stage in the analysis and examine whether at the succeeding visits, positive GI cancers tend to be diagnosed at late stages/more aggressive grades indeed.

5. One of the developed model assumptions in this study was that detecting GI cancer early depends on minimizing the delay time. However, considering the former point—the possibility for more aggressive GI cancer to be initiated in the time interval between two referrals—leads to the conclusion that detecting GI cancer early depends also on the frequency of the investigations. The effect of

investigations frequency on the transition rates should be assessed as well.

6. Though a normal progressive disease model will end up with “death” state, death state was not included in the developed model. Adding death state to the model could help to examine the over-diagnosis of nonprogressive or very slow-growing GI cancers.

7. Most importantly, in the developed model, “incomplete investigation” state was presumed as a mutually exclusive state from positive and negative GI cancer states, as only the “observed” states in the patients’ admission history were considered. However, a patient who is observed in the state of incomplete investigation might be healthy (negative GI cancer) or have a hidden GI cancer that can be diagnosed by clinical investigations. Accordingly, to incorporate this possible misclassification, a hidden Markov model should be fitted to distinguish between the observed states and the truly underlying states of the IDA patients’ admissions as proposed in the diagram [Figure 6-5].

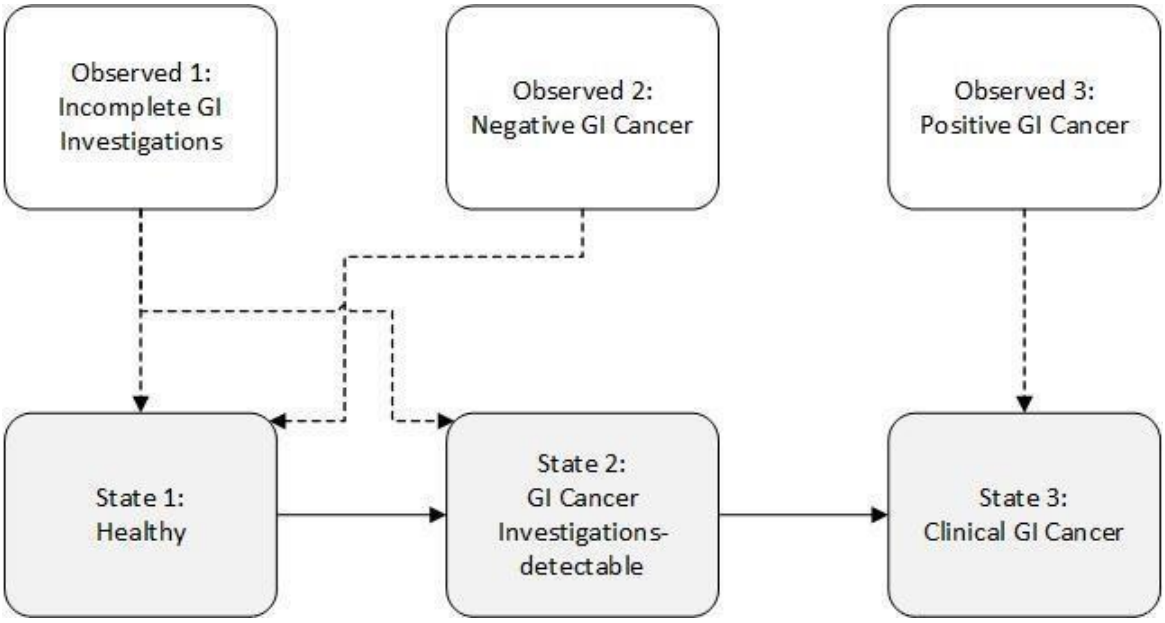


Figure 6-5 Markov three-state diagram: the white boxes are the observed states and the gray boxes are the true underlying state. The solid lines are the transitions between true states. Observation of an incomplete investigations could be truly healthy or misclassification of an investigations-detectable GI cancer

The strengths of this study are that it represents the first study that demonstrates the appropriate methods to model the IDA patients' episodes of care at a secondary-care center. It also raises the awareness of the importance of completing the GI investigations especially in IDA patients who are at high risk of GI cancer and fit enough to do the investigations-. The estimation of the transition rates and length of delay time in the state of incomplete investigations in future large-studies can help policy makers to establish what is the maximum delay time a confirmed IDA patient should not be allowed to stay in before investigated, and what are the measures that could be put in place to reduce or minimize this time.

Acknowledgement

The author would like to thank Dr Jonathon Snook, Dr Sharon Docherty, and Professor Peter Thomas for their advice on this study.

Ethical approval

Retrospective analysis of anonymised secondary data, external ethics approval was not required. Bournemouth University ethics approval was attained on 22/02/2018, reference id: 19925.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Almilaji O, Parry S, Docherty S, Snook J. Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia. Sci Rep 2021;11:13055.
2. Hung N, Shen CC, Hu YW, Hu LY, Yeh CM, Teng CJ, *et al*. Risk of cancer in patients with iron deficiency anemia: A nationwide population-based study. PLoS One 2015;10:e0119647.
3. Silva AC, Sheppard ZA, Surgenor SL, Williams EJ, Thomas PW, Snook JA. Clinical risk factors for underlying gastrointestinal malignancy in iron deficiency anaemia: The IDIOM study. Frontline Gastroenterol 2014;5:237-42.
4. Stone H, Almilaji O, John C, Smith C, Surgenor SL, Ayres L, *et al*. The dedicated iron deficiency anaemia clinic: A 15-year experience. Frontline Gastroenterol 2020. doi: 10.1136/flgastro-2020-101470.
5. Almilaji O, Smith C, Surgenor S, Clegg A, Williams E, Thomas P, *et al*. Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia. BMJ Open Gastroenterol 2020;7:e000403.
6. Wijayasekara CM, Surgenor SL, Hatcher CA, Clegg AJ, Williams EJ, Snook JA. Clinical risk factors for underlying gastrointestinal malignancy in iron deficiency anaemia-prospective validation of the IDIOM score. Cogent Med 2016;3:1213216.
7. Surgenor SL, Kirkham S, Parry SD, Williams EJ, Snook JA. 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, Scott BB; British Society of Gastroenterology. Guidelines for the management of iron deficiency anaemia. Gut 2011;60:1309-16.
9. Rockey DC. Gastrointestinal tract evaluation in patients with iron deficiency anemia. Semin Gastrointest Dis 1999;10:53-64.

10. Liu K, Kaffes AJ. Iron deficiency anaemia: A review of diagnosis, investigation and management. Eur J Gastroenterol Hepatol 2012;24:109-16.
11. Zhu A, Kaneshiro M, Kaunitz JD. Evaluation and treatment of iron deficiency anemia: A gastroenterological perspective. Dig Dis Sci 2010;55:548-59.
12. Rai S, Hemingway D. Iron deficiency anaemia: Useful diagnostic tool for right sided colon cancers? Colorectal Dis 2005;7:588-90.
13. National Institute for Health and Care Excellence (NICE). Suspected Cancer: Recognition and Referral 2015, updated Sep 2020. Available from: https://www.nice.org.uk/guidance/ng12. [Last accessed on 2020 Nov 08].
14. Snook J, Bhala N, Beales ILP, *et al*. British Society of Gastroenterology guidelines for the management of iron deficiency anaemia in adults. Gut 2021. doi: 10.1136/gutjnl-2021-325210.
15. Almilaji O, Snook J, Thomas P. The IDIOM App (ver 1.0): A Web-based Application to Predict the Risk of Gastrointestinal Cancer in Iron Deficiency Anaemia. [Software Medical Device, class I, MDD]. UK: BU Innovations Limited (BUI); 2020 Available from: https://www.predict-gi-risk-in-ida.com. [Last accessed on 2021 Apr 06].
16. Welton NJ, Ades AE. Estimation of Markov chain transition probabilities and rates from fully and partially observed data: Uncertainty propagation, evidence synthesis, and model calibration. Med Decis Making 2005;25:633-45.
17. Briggs A, Sculpher M. An introduction to Markov modelling for economic evaluation. Pharmacoeconomics 1998;13:397-409.
18. Shannon CE. A mathematical theory of communication. Bell Syst Tech J 1948;27:379-423.
19. Sonnenberg FA, Beck JR. Markov models in medical decision making: A practical guide. Med Decis Making 1993;13:322-38.
20. Grüger J, Kay R, Schumacher M. The validity of inferences based on incomplete observations in disease state models. Biometrics 1991;47:595-605.

21. Von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies.

Chapter 7 : The development of a clinical decision-support web-based tool for predicting the risk of gastrointestinal cancer in iron deficiency anaemia; the IDIOM App

Open access original research article published at: https://www.mdpi.com/2673-6470/2/1/7

Citation: Almilaji, O., Engen, V., Snook, J., Docherty, S. The Development of a Clinical Decision-Support Web-Based Tool for Predicting the Risk of Gastrointestinal Cancer in Iron Deficiency Anaemia—The IDIOM App. 2022. *Digital (Special Issue Intelligent Digital Health Interventions)*; 2, 104-119. https://doi.org/10.3390/digital2010007

Abstract

To facilitate the clinical use of an algorithm for predicting the risk of gastrointestinal malignancy in iron deficiency anaemia – the IDIOM score, a software application has been developed, with a view to providing free and simple access to healthcare professionals in the UK. A detailed requirements analysis for intended users of the application revealed the need for an automated decision-support tool in which anonymised, individual, patient data is entered and gastrointestinal cancer risk is calculated and displayed immediately, which lends itself to use in busy clinical settings. Human-centred design was employed to develop the solution, focusing on the users and their needs, whilst ensuring that they are provided with sufficient details to appropriately interpret the risk score. *The IDIOM App* has been developed using R Shiny as a web-based application enabling access from different platforms with updates that can be carried out centrally through the host server. The application has been evaluated through literature search, internal/external validation, code testing, risk analysis, and usability assessments. Legal notices, contact system with research and maintenance teams, and all the supportive information for the application such as description of the population, intended users have been embedded within the application interface. With the purpose of providing a guide of developing standalone software medical devices in academic setting, this paper aims to present the theoretical and practical aspects of developing, writing technical documentation, and certifying standalone software medical devices using the case of the *IDIOM App* as an example.

Keywords: software medical device; digital health; medical risk prediction models, decision support systems, gastrointestinal cancer, GI, iron deficiency anaemia, IDA.

1. Introduction

The association between iron deficiency anaemia (IDA) and gastrointestinal (GI) cancer is well documented in the medical literature [1-6]. As a consequence, IDA finding in at-risk patients is considered as reason for suspected cancer referral to secondary care in the UK [7]. To predict the GI cancer risk in IDA, and stratify the patients in risk groups accordingly, we have previously built a binary multivariable logistic model to predict the risk of GI cancer in patients with confirmed IDA– the IDIOM score; **I**ron **D**eficiency as an **I**ndicator **O**f **M**alignancy [4], based on four simple variables: age, sex, haemoglobin concentration (Hb), and mean cell volume (MCV).

With a view to providing free and open access to healthcare professionals in the UK, a digital decision support tool called the *IDIOM App* [8], was developed in 2018-2020. The app consists of two major components. These are:

- An internal algorithm to predict the GI cancer risk in IDA patients, IDIOM score.
- A web−based interface to allow selecting input data (sex, age, Hb, MCV); presenting the results (prediction of the GI cancer risk estimate, and the risk group); the app terms of use, privacy and cookies policy, and app description; enabling the communication with the app research and maintenance teams; and printing (or saving) the results when needed.

The *IDIOM App's* model was based on retrospective examination of secondary clinical datasets for adult patients (n=2390) referred to Poole Hospital IDA clinic between 2004 and 2018. The anonymized dataset includes age at presentation, sex, blood test results (Hb, MCV, and iron studies (Ferritin and /or Transferrin saturation)), and the diagnostic findings on standard investigation of the upper and lower GI tract (i.e. the presence or absence of GI cancer).

The model did not include other potential risk factors of GI cancer such as family history, previous cancer, race, unintentional weight loss, red meat consumption, alcohol, and tobacco use, etc. Therefore, it is not possible to

use the *IDIOM App* to diagnose GI cancer, and its results must be viewed as reference information only and interpreted in the context of each patient's individual circumstances. Whilst using the app, healthcare professionals should rely on their own independent clinical judgement and knowledge, relevant referral guidelines, and on assessing all the GI cancer risk factors including the factors which have not been considered in the app when taking any decision relating to the patient's future clinical management.

There are, currently, limited resources on developing standalone software medical devices. Therefore, this paper aims to present the theoretical and practical aspects of developing, writing technical documentation, and certifying standalone software medical devices using the case of the *IDIOM App* as an example. Unlike the structure of typical research papers; introduction, methods, results, and discussion (IMRAD), the structure of this paper starts with the introduction, then follows the logical order of creating and putting the App into service, and concludes with a general discussion of lessons learnt and considerations for interested readers.

2. Compliance to standards

The first step in any medical device development project is to know what the applicable laws to the medical device are. Like any other set of regulations, these laws depend on time and place. Time is a factor of which the law takes cognizance, and it affects the laws in many issues such as commencing date, terminating date, legal duration, and the retrospective effect of legislation [9]. Place refers to the locations of performance and jurisdiction, *i.e.* the location of the manufacture, and the location in which the medical device will be placed on market.

Standards are considered as the minimum regulatory requirements medical devices should be satisfying. In general, standards are documents written by national or international committees to document the "state of the art". Examples of these standards in the EU are Medical Device Directive (MDD)

[10], and Medical Device Regulations (MDR) [11]. In addition, manufacturers of medical devices are advised to develop their products in adherence to harmonized standards (depending on the nature of the device) such as IEC 62304 [12], IEC 62366 [13], ISO 14971 [14], ISO 9241-210 [15], ISO 13485 [16] and other relevant guideline documents [17-25]. Although these harmonized standards and guideline documents are not legally binding in the EU, demonstrating that software medical devices have met the legal requirements would be difficult without them.

The IDIOM App was developed in adherence to the EU-wide MDD [10], which has recently been superseded by the MDR [11] . Since the MDR has come into force in 26 May 2021 this means medical devices which lawfully placed on the EU market pursuant to the MDD prior to 26 May 2021, may continue to be made available on the market up to 5 years from the certificate's issue/renewal date or 4 years from the MDR date of application, whichever comes first. The MDD, and all relevant harmonization standards/guideline documents to software medical devices were followed whilst developing the *IDIOM App*.

3. App risk classification

The second step is to confirm that the software is a medical device, and if so what risk class it is. Because the *IDIOM App* software combines medical knowledge databases and algorithms with patient specific data; it is considered as a "decision support software" [21]. The *IDIOM App* does not allow direct diagnosis of the gastrointestinal cancer by itself and only provides reference information to enable healthcare professionals to make clinical decisions as they ultimately rely on their own knowledge. However, given that any "decision support software" that applies automated reasoning i.e., a prediction algorithm in which the healthcare professional does not review the source/raw data, may be considered as a medical device that falls within the scope of the MDD; the *IDIOM App* was considered as a medical device.

Because this software works alone and not in combination with any physical medical device; it is a “stand-alone software” [21] or so called “software as a medical device”(SAMD) [23]. Medical device classification rules are based on the impact of the device on patients or users and the potential risks associated with the technical design and production of the devices [10]. As no direct diagnosis for GI cancer is possible based on the information provided by the *IDIOM App*, nor is this application diagnosing a vital physiological process, rule 12 from section 3, III Classification annex IX, of the MDD may be applied and the app can be classified as Class I. Since the *IDIOM App* is a stand-alone medical device software; it is an “active device” according to chapter one (I definitions), annex IX, in the MDD [10]. Consequently, the full and final app classification was “stand−alone, clinical decision−support (CDS) software, none−sterile, none−measuring, none−reusable surgical instrument, active Class I medical device”.

According to the MDD, for any app that falls within the definition of a medical device to be put into service and be used by health professionals in the European Community, it must bear the *Conformité Européenne* (CE) mark. CE mark cannot be affixed to any medical device app unless the app is registered with the competent national authority in the country it would be put in service at. In the UK, all apps must be registered with the UK Medicines and Healthcare Products Regulatory Agency (MHRA). Registering the app with the MHRA cannot be done unless the app indicates its conformity with the provisions of the MDD through signing the declaration of conformity (DoC). Indicating the conformity with the provisions of the MDD is carried out through following the appropriate conformity assessment procedure.

After Brexit, CE marking will continue to be recognised in Great Britain until 30 June 2023. As of 1 July 2023, a UKCA (UK Conformity Assessed) mark will be required in order to place a device on the Great Britain market. Until 30 June 2023, manufacturers can use the UKCA mark on a voluntary basis. From 1 July 2023, a UKCA mark will be needed in order to place a device on the Great Britain market. However, Class I devices that have not a measuring

function nor are sterile can be self-certified against the UKCA mark. The Great Britain route to market and UKCA marking requirements is still based on the requirements originated from current EU legislation [26].

4. Relevant conformity assessment

The third step is to identify the relevant conformity assessment route for the device. Conformity assessment procedures differ according to the classification of the medical devices. For medical device Class I, like the IDIOM App, Annex VII in the MDD must be followed to draw up the DoC before placing the medical device on the market, and to prepare the technical documentation that allow the assessment of the conformity of the product. The details of what to include in each part of this technical documentation and who should be responsible of assessing it depends on the description and classification of the app, the relevant regulations, and harmonized standards.

So, for instance:

- While high-risk software medical devices need to include the documentation for every development process such as design, integration, and testing, according to EN 62304, only development planning, requirements analysis, implementation and release are needed to be included in the documentation for the development process of Class I software medical devices according to the same EN 62304 standard.
- “instructions for use” which are required by the MDD for higher-risk medical devices are not required for Class I medical devices.
- Descriptions of “used methods and validation report” which are required by the MDD for Class I medical devices that are placed on the market in a sterile condition, are not applicable for software Class I medical devices.
- While compliance of Class I devices is based on self-declaration, all other higher-risk devices require use of an approved notified body to assess compliance.

The following figure illustrates the needed steps to put software medical device on market:

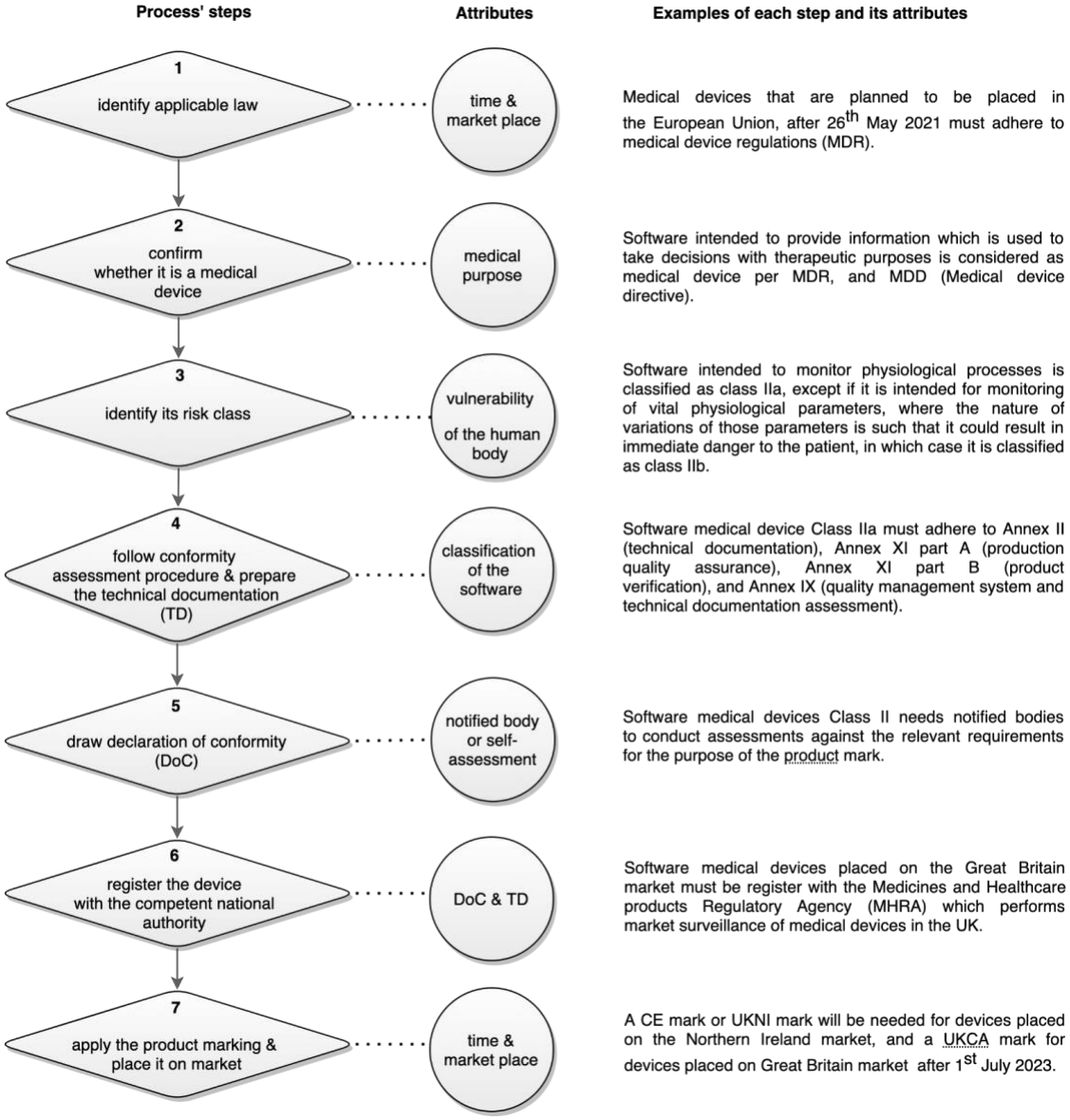


Figure 7-1 Steps of placing a software medical device on the market

After a self-assessment for conformity certification was conducted to apply the CE marking, the technical documentation of the *IDIOM App*, is established per Annex VII in the MDD to include: the app general description and its intended use(s); development planning; requirements analysis; implementation, and deployment; clinical evaluation and interface usability assessment; risk analysis, maintenance and plan for post-market surveillance; and release and label. Also, the code, the data, the signed

declaration of conformity, and a list of all the harmonization legislations and standards that has been adhered to during the app development and the writing of the technical documentation were included in the technical documentation. A version control copy of the technical documentation is kept and updated by the App' research and maintenance team at the University.

5. Technical documentation

5.1 App general description

5.1.1 What the App does?

Using four predictors (input data); sex, age, Hb, and MCV, the *IDIOM App* calculates the risk of any type of GI cancer for a specific iron−deficient patient. The results of the calculations are displayed in a table that contains the selected predictors' values of sex, age, Hb, MCV for the patient, the risk estimate with its 95% confidence interval and the risk group of the patient based on the risk estimate. This table is followed by an explanation of the risk estimate, confidence interval of the risk, and the risk groups. The risk estimate represents a probability (in a percentage format) that an individual confirmed ID patient with the particular set of predictors entered will prove on investigation to have cancer somewhere in his/her GI tract. Though this probability risk provides a realistic estimate of any potential GI cancer, it implies no certainty about the presence of GI cancer. The confidence interval of the predicted risk represents a range of values that predicts where the risk will fall for a population of confirmed ID patients who share the selected values of sex, age, Hb, and MCV with 95% confidence interval. Risk groups are classifications that describe IDA patients who fall within certain ranges of risk estimates values of positive GI malignancy.

- If the predicted risk of GI cancer is very low, the risk figure will be displayed in dark green font colour and the risk and groups' cells in light green background colour.

- If the predicted risk of GI cancer is low, the risk figure will be displayed in dark green font colour and the risk and groups' cells in white background colour.
- If the predicted risk of GI cancer is moderate, the risk figure will be displayed in black font colour and the risk and groups' cells in white background colour.
- If the predicted risk of GI cancer is high, the risk figure will be displayed in red font colour and the risk and groups' cells in white background colour.
- If the predicted risk of GI cancer is very high, the risk figure will be displayed in red font colour and the risk and groups' cells in amber background colour.

A screenshot of the *IDIOM App* is shown in Figure 7-2, in which a very low Gi cancer risk is predicted.

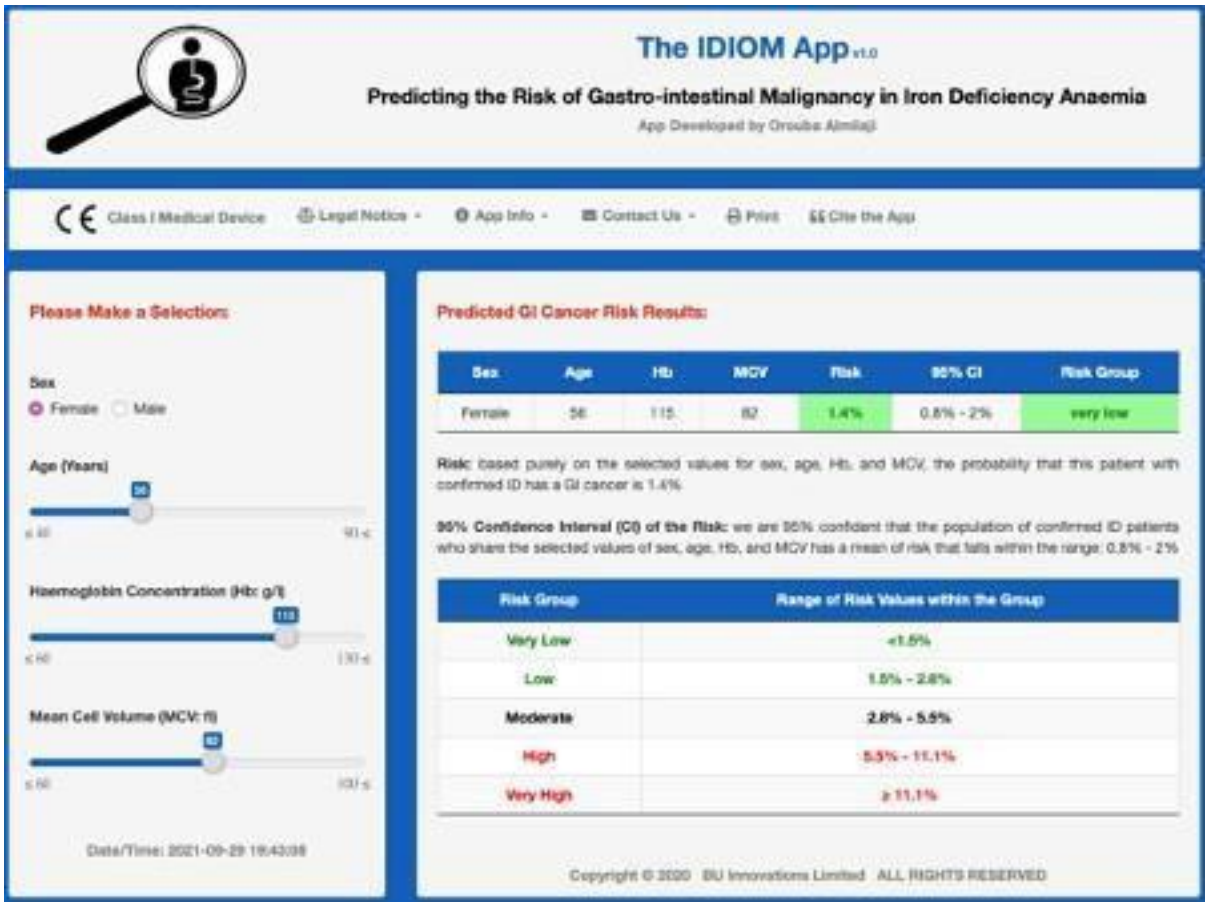


Figure 7-2 A screenshot of the IDIOM App

5.1.2 Who is the App for?

The intended patients' population for this app are the confirmed iron−deficient patients (adults only). Confirmed iron deficiency is defined by standard laboratory criteria; transferrin saturation <15% and / or serum ferritin concentration less than the lower limit of the reference interval for the laboratory. The app is not intended to be used on patients who have been given iron replacement therapy prior to their blood testing as their Hb and MCV might be elevated by this therapy and might cause the app's calculation, which relies on these blood markers values (Hb, MCV), to be unreliable.

5.1.3 Who should be using the App and where?

The *IDIOM App* is designed as an adjunct to standard counselling and personal discussion with a healthcare professional and cannot replace it. The intended targeted end users for the software are healthcare professionals only such as gastroenterologists, and specialist nurses. The intended environment in which this app should be used is a clinical setting such as IDA clinics, within gastroenterology departments in hospitals.

5.1.4 Manufacture

Because the app was a research outcome of a match−fund PhD project between Bournemouth University (BU) and University Hospitals Dorset NHS Foundation (Poole Hospital); At the conclusion of the project, BU assigned the IP to BU Innovations Limited (BUI) for the latter to be the designated legal owner and manufacturer of the current version of the App.

5.2 Development planning

The development of the *IDIOM App* was conducted using an agile approach in which a loop of different tasks is completed through multiple iterations. Every iteration aimed to deliver target milestones by working on described tasks and had scheduled start and end dates. Tasks include planning,

requirement analysis, design, implementation, testing, and release illustrated in Figure 7-3.

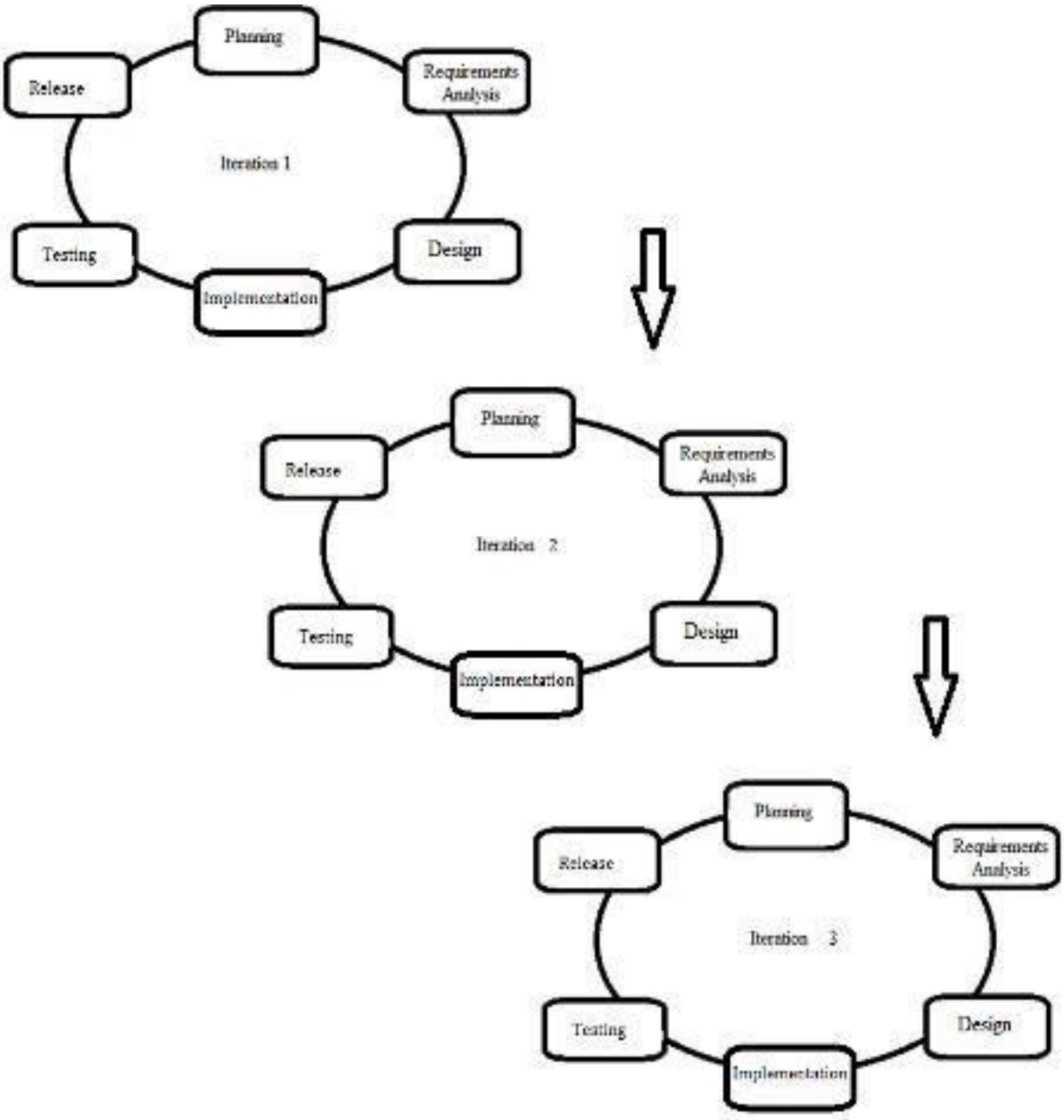


Figure 7-3 The agile model for developing the IDIOM App with three iterations, each iteration includes six tasks: planning, requirement analysis, design, implementation, testing, and release

As an example, for the “design” task, target milestones include function hierarchy diagram, screen layout diagrams, pseudo code, entity-relationship diagram with a full data dictionary, etc. For each iteration, planned number of these milestones are accomplished.

5.3 Requirement analysis

The requirements of the intended users of the *IDIOM App* were gathered and understood through:

- Interacting directly with the expected end users by working as an honorary research fellow in a gastroenterology department which has its own dedicated IDA clinic [3].
- Showcasing the early version of the app in presentations at conferences and gastrointestinal departmental clinical governance meetings. These opportunities enabled the interaction with the experts in the domain, learn more about the rational of the app, and how to improve it [27-32].
- Other similar apps such as 'predict prostate app' [33] that has been developed in Cambridge University to provide cancer-specific and overall percentage survival estimates for up to 15 years, were another very useful source to envisage probable solutions.

The requirements were documented in a requirement specification table, according to their type, the description of each requirement, and the proposed solution for each requirement. Specified requirements include: licensing, registration, terms of use, data protection and privacy notice, release notice, security, printing, ease of learning, understandability, structure & visibility, subjective satisfaction, end users' feedback and enquiries, accuracy, availability, safety, response, installation, adaptability, compatibility, level of support, and maintainability. Requirements main types were functional, legal, usability, reliability, performance, and supportability.

The prioritization of the requirements is carried out according to Moscow method [34] in which each requirement was categorized as (must have, should have, could have, or won't have). At the last stage of requirements analysis, an ordered list of interactions between the actors and the app has been illustrated through a "Use Case".

5.4 Implementation

A web-based design was chosen for the app after considering different issues such maintenance, compatibility, security, speed and performance, and overall control. The programming language which was used develop the app is R. Because it is free, highly extensible, and it was the same language used to run the statistical analysis. Two additional R packages were used when building the app: namely Shiny package and DT package. However, since R (and its packages) is an open source language, the general public license (GPL) must be considered carefully if “commercializing” is an aim for other medical devices. Using Shiny Package, two R scripts have been created in the app’ folder; a user interface object (ui.R) and a server function (server.R).

The interface controls the layout, appearance, and facilities entering the users’ inputs and displaying the outputs of the model. The server contains the instructions that needs to run the app. The app’s folder contains all the resources required to build the application such as the prediction model (which was saved as *R* object without the data), logo, and support information *HTML* pages.

5.4.1 The Interface

Since the *IDIOM App* end users who are healthcare professional who work usually in very busy environment, the user interface was designed and implemented to avoid any computing/statistics jargon language, and unnecessary explanations such as what is meant by Hb or MCV. The number of the main panels in the display area was limited/minimized. And the navigation between them made predictable by following a natural reading pattern i.e. the English language reading pattern (the title and subtitle come at the top, then the direction is left to right). Screen layout diagrams and the specifications for the app interface were documented. For each artefact, a list of its attributes and special design considerations are described. For the navigation panel, attributes include, CE marking, legal notice menu, app info menu, contact menu, print (or save) command, and cite the app box message.

Considerations include placing appropriate icons in front of each former panel element to easily identify them, are described in detail. A responsive user interface has been implemented by calling the function *fluidpage()*. *Fluidpage()* easily adapts and responds very well on all devices (desktop, laptop, tablet, mobile, etc.). Even when changing the orientation of the mobile device from landscape to portrait mode, the design would change automatically without reducing the visible content. The collapsible feature was used also in the implementation of the app tables. So, when the result table cannot be fitted in a small browser window, the table will be collapsed to fit the width of the screen.

5.4.2 The IDIOM App logo

The logo has been created by the same app developer -the PhD candidate; O.A.M- and designed to be as simple as possible by using only two colours (black and white), and by depicting a human gastrointestinal tract inside a magnifier to reflect the fact that this app is examining something relating to the GI tract.

5.4.3 Text, font, and colours

To look more friendly and less complex, the displayed numeric values in the text have been rounded to have a maximum of two digits only. The font size was selected to be readable from an arm length distance. Also, the colour scheme and font choices which have been selected to be used in the app were as consistent as possible to NHS identity colours and fonts. This is because the users of the app are mostly health professionals who work within the NHS. Thus, the font families which have been used in the app were Frutiger and Arial. Frutiger is a clear and easy to read at a distance and in small sizes. As the colours blue and white are strongly associated with the NHS with. The NHS Blue is the dominant colour in the app colour palette. Because red colour is typically used to refer to danger or emergency, it has been used and one of its shades (amber) to refer to the positive GI malignancy

text and cell background. While the green colour has been used to refer to the negative GI malignancy.

5.4.4 Mouse/pointing devices

No keyboard or sound effects have been used. Only mouse and pointing devices are used and a single click (or tap on touch-screen devices) is enough to select a value e.g. the sex variable.

5.4.5 Server function

The server-side is defined to accept inputs and compute outputs by assigning reactive expressions to output slots. Reactive expressions cache their values and know when their values have become outdated. This means at that the first time when a reactive expression runs; it will save its result. So, the next time the reactive expression is called, it can return this saved result without doing any computation (which will make the app faster). *Renderdatatable()* is used to generate output. This reactive wrapper returns special expressions that are only re-executed when their dependencies change. This behaviour is enabled the app to automatically update output whenever input changes. A specification for the app's the function hierarchy was documented, in which each function is listed along with its initiator (executed by), and the steps of running this function (executed through). So, for instance, the "select patient' data" function, will be executed by the end user through ticking the right sex input box, and moving the inputs sliders in the inputs panel to choose the values of age, Hb, and MCV.

5.5 Deployment

The *IDIOM App* was deployed online by setting up a single instance of the app on a server connected to the internet. The chosen online platform to deploy the *IDIOM App* was a cloud server. Before uploading the app to this virtual server, the server was configured by setting up a secure access and firewall, installing *R* & *DT* / *Shiny* packages from the comprehensive R archive

network (CRAN), installing *nginx* web server in order for the app content to be visible to the public through Hypertext Transfer Protocol (HTTP), and adding Secure Sockets Layer (SSL) certificate to the HTTP to get HTTPS. After that, and to make the Uniform Resource Locator (URL) address of the app an easy address to remember, a dedicated domain called predict-gi-risk-in-IDA.com has been created and purchased. Finally, the *IDIOM App* website was registered with Google search console and optimized for search engines. No personal or identifying information of users that are accessing or using the app, are gathered by the *IDIOM App*.

5.6 Clinical evaluation

To clinically evaluate the *IDIOM App*, a valid clinical association between GI cancer and IDA was established through existing evidence literature searches. In fact, five previous studies have examined the association between gastrointestinal cancer and iron deficiency anaemia and developed a multivariable risk algorithm to predict the risk of GI cancer in IDA (35, 36,37,38, 2]. The sample size for these studies was 98, 148, 695, 643, 720 respectively. Though age was a universal positive predictor of the GI cancer risk, as expected, in all these studies, the results were conflicting with regard to the other predictors; sex, Hb, iron studies, and MCV.

One explanation for these inconsistent results might be caused by the small size of the studies especially in Capurso et al. 2004 and Ho et al. 2005. [35-36]. Another explanation could be the forcing of the quarters or dichotomous classification of continuous predictor variables in the predictive model. Age and Hb for instance, were coded into categories in Silva et al. 2014 study [2] and James et al. 2005 [37] study. Categorization of continuous data should be avoided in the statistical analysis as it leads to information loss, underestimation of the extent of variation in outcome between groups, and concealment for any non-linearity in the relation between the variable and outcome [39]. Then an analytical validation for this association was carried out by using previously collected patients' data (n=1879) who were assessed

between 2004 and 2016 inclusive at the Poole IDA clinic [4]. Finally, a clinical validation was carried out by assessing the app model performance using internal and external clinical datasets.

The IDIOM model was internally validated using an anonymized clinical dataset from Dorset [4], and externally validated using two anonymized clinical datasets from Oxford and Sheffield [40]. The criteria for inclusion in all the datasets were iron deficiency confirmed by standard laboratory criteria, and subsequent investigation of the upper and lower GI tract. The internal Dorset dataset was were collected for the period 2017-2018 and comprised in total 511 subjects with IDA referred to a dedicated IDA Clinic. The Oxford dataset was collected for the period 2016-2019 and comprised 1147 subjects with IDA referred for fast-track investigation. The Sheffield dataset was collected for the period 2013-2018 and compromised 477 subjects with IDA referred to a dedicated IDA Clinic. The training and internal datasets were merged to form the Dorset dataset (2390=1879+511) which was used to fit the updated full IDIOM model. After this, the full IDIOM model was regulated using Lasso method (least absolute shrinkage and selection operator). The final updated regulated multiple binary logistic regression of the IDIOM model was constructed according to the formula [40]:

$$\log\left\{\frac{\mathbb{P}(GI\ Maligancy = postive)}{\mathbb{P}(GI\ Maligancy = negative)}\right\} = -1.84 + 0.89\ sex + 0.05\ age - 0.03\ MCV - 0.06\ Hb$$

There were differences between the datasets, as shown in table 7-1. As expected, the Oxford dataset had a lower median Hb in particular, as subjects presented exclusively through the fast-track pathway.

Table 7-1 Descriptive statistics for the three datasets

Dataset		Dorset	Oxford	Sheffield
Dataset size	N	2390	1117	474
GI Cancer	Positive - n (%)	200 (8.4%)	86 (7.7%)	36 (7.6%)
Sex Ratio	M/F	0.56	0.66	0.92
Age (Years)	Median (Q1, Q3)	71 (59, 79)	74 (65, 81)	69 (61, 77)
Hb (g/l)	Median (Q1, Q3)	104 (93, 113)	91 (79, 101)	104 (95, 116)
MCV (fl)	Median (Q1, Q3)	80 (74, 86)	81 (75, 87)	80 (76, 86)

Statistical The goodness of fit for the IDIOM model was satisfactory (by examining the deviance and residual test, smoothed scatter plot, variance inflation factor, Cook's distance and standardised residual errors, analysis of variance $\chi 2$ test, Akaike information criterion, pseudo R^2) [35].

By estimating measures of discrimination, calibration, and clinical utility using the external validation datasets, the predictive performance of the app's model was assessed. The discrimination of the IDIOM model using the external validation data was 70% (95% CI 65, 75) and 70% (95% CI 61, 79) for the Oxford and Sheffield datasets respectively. The analysis of calibration showed no tendency for under or over-estimated risks in the external validation datasets. Decision curve analysis showed the clinical value of the model with a net benefit that is higher than 'investigate no-one' and 'investigate all' strategies up to a threshold of 18% in the external validation datasets. Using a risk threshold of around 1.2% to categorise patients into the ultra-low risk group showed that none of the patients stratified in this risk group proved to have GI cancer on investigation in the training, internal, and external validation datasets. Therefore, the validation has demonstrated promising results for the IDIOM model in predicting the risk of underlying GI cancer in independent datasets collected in different clinical settings [35].
Further work is planned to compare the IDIOM model performance, which is built using logistic regression, to other machine learning methods such as

random forest and support vector machine and multi-layer perceptrons (a type of artificial neural network commonly used for structured, numeric, data). Conceptually these types of machine learning models work the same, but may yield further performance improvements.

5.7 Interface usability assessment

To evaluate usability of the app interface, standard usability questionnaire applied. Participants include NHS staff such as IDA nurse specialists, gastroenterologists, etc. Participation was voluntary and anonymous. Participation was done at two points of time, where at the first time point, four participants assessed the interface usability. And at the second time point, three participants did the same. Two participants among those who assessed the usability at the second point of time, were also among those who assessed the app at the first point of time. Each participant tried using the app then commented on the:

- Ease of use, in terms of keystroke level model (KLM) [41].
- Understandability, in terms of what the app does, its intended use, etc.
- Structure and visibility including app's layout, font, familiarity, interface elements, clarity, navigation through the main panels in the interface, colours, readability, etc.

After commenting on the former aspects of the app's interface, each user has given an overall satisfaction score (scale 1 to 10, in which the higher the number the better is the interface), and provided an open feedback. Feedback, generally, were revolved around changing the explanations of the risk estimate and confidence interval to a more lay English. All users' feedback was taken on board and the interface was changed accordingly. Usability assessments (n=7) have shown a promising overall mean user satisfaction score of 8.5 out of 10. Notably, mean user satisfaction score was higher at the second point of time.

5.8 Risk analysis, maintenance and plan for post-market surveillance (PMS)

Risk management techniques were applied throughout the life cycle of the *IDIOM App*. The risk analysis includes the documentations for:

- risk management plan,
- initial hazard identification and risk assessment,
- risk control,
- and evaluation of the overall residual risk.

The risk management plan includes responsibilities, risk review requirements, risk acceptability levels, reference to standards, verification activities, criteria for risk acceptability, overall residual risk acceptability, production and post-market activities.

Initial hazard identification and risk assessment includes a description of the intended use(s)/purpose(s) of the device, the intended patient population, the users and the use environment, a list of all qualitative or quantitative characteristics that could affect safety, known or foreseeable hazards that are/could associated with the device in normal and fault conditions, causes, consequences, and associated risks identifications in terms of severity and probability of the harm occurring.

All the identified hazards for the *IDIOM App* were low/acceptable foreseeable risks. Examples of these risks are:

1. Denial-of-service (DoS). This risk might be caused by external technical attack.
2. Using the app to predict GI cancer for the wrong population. This risk might be caused by human error.

The risk control includes for each risk, the risk type, description of the risk identified, risk, elimination/reduction measures, evaluation of the risk at the reduced level, probability and severity of the risk.

Examples of measures applied to mitigate the risk of “using the app to predict GI cancer for the wrong population” include providing a clear definition of the

appropriate patient population in the welcome page, the description page of the app, and in the terms of use.

The evaluation of overall residual risk includes a list of all the residual risks that were identified in the risk control, with their new status after risk controls have been applied, evaluation of the overall residual risk, additional control measures need to be applied, and justification of the updated status of the overall risk.

All the residual risks of the app were acceptable apart from one low-risk (the DoS risk). Yet, the overall residual risk was acceptable because any further reduction was impractical (not possible) to DoS risk. Since the app is web-based application, external DoS risk cannot be 100% prevented.
Plans for post-market maintenance, and reactive/proactive surveillance have been established. Maintenance is expected to be run routinely during the expected lifetime of this app, and it will involve bug−fixing and routine updates.

5.9 Release and label

All the labelling (also be referred to as “information supplied”) content have been provided in an electronically accessible forms that can be accessed directly through the app webpage and were subject to document (version) control principles. The *IDIOM App* labelling was delivered in human readable format and included: app description, terms of use, R’ GPL, privacy and cookies policy, research teams details, research team publications, app title, CE Marking sign, date/time of access, app logo, BUI copyright notice, and the developer’s name.

6. Discussion

Though the *IDIOM* App was classified as low-risk by the MDD, placing it on market was a time-consuming process, as it was new to several parties involved. After three years of full-dedicated time, the process of developing, and writing the technical documentation was completed and the *IDIOM App* Version 1.0, was successfully registered with the MHRA and lawfully placed on the market pursuant to the MDD on 1st Dec 2020, with expected service life up to June 30th 2023 under the present certification.

During this process, many lessons were learned, and many hurdles were overcome, these involved:

1. Most and foremost, being a de novo software medical device development project in an academic setting has demanded a lot of initiative extra learning by the research team as the university did not have the pathway and processes that supported digital medical device development when the project has started. Nevertheless, without the genuine willingness of BU to do things differently such as setting up surveillance system when the app went live, the *IDIOM app* was not distanced to succeed. BU is now developing quality managed device development process as a consequence of this project success.

2. Being part of a PhD project, there were no available resources to outsourcing the app development, developing the app and producing the technical file that met certification standards categorically demanded time-consuming fastidious attention to details and record-keeping. This is not a light undertaking and is actually very tough without an existing framework.

3. Coordinating with many stakeholders with different perspectives and priorities within the university (research development and support department, legal department, IT service, etc.), the trust, and external

consultancy which has required a high level of commitment, negotiation skills, and clear communication.

4. Finishing the project within a limited PhD timeframe that caught up in the stressful Covid-19 pandemic period. A period which did not only witness the delaying of the enforcement of the new EU regulation (MDR) but also entangled the uncertainty about the new applicable medical device laws in the UK after the Brexit. Changing the applicable laws, actually, makes the case of *IDIOM App* project a very interesting case.

The app is currently used in the UK secondary care such as Poole hospital as decision support tool. Recently, it was endorsed by the British Society of Gastroenterology guidelines for the management of iron deficiency anaemia in adults [42]. The total up-to-date number of the app users is 762. According to Google analytics, these users are from the UK (78%), Spain (4%), India (2%), and from 45 countries around the world. Most of the access to the app comes through desktop (76%), then mobile (23%), and finally tablet machines (1%).. And the most visited pages are the app page itself, the publications page, and research team page.

Experience gained through developing, updating, and interacting with the app's prospective users through the embedded contact system in the app might help to develop the app further. Conditional on being successful in future funding, further plans are proposed for the new version of the app to include:

- Certifying the app pursuant to the new UKCA and MDR regulations. The MDR regulations are more strict than the MDD, and the *IDIOM App* might be confirmed by the MDR as a medical device and not as a borderline. This is because "disease prediction" is included now as a new *medical purpose* in the new MDR regulations.
- Using the App in primary care, subject to clinical validation, as a decision-support tool to refer patients to secondary care.

- Expanding the sample size, and adding new variables such as family history, BMI, and the FIT test to the prediction model after examining their predictive values.
- Validating the app externally on new clinical datasets for patients from outside the UK.

The strength of this study it represents the first study to document the development of a standalone software medical device in an academic setting from a practical experience. And to discusses the hands-on aspects and hurdles that academics may face when developing medical devices. The limitations include the fact that the process of developing and certifying the *IDIOM App* was done according to the MDD regulations which were subsequently supressed by other regulations. Yet, the process of developing standalone software medical device still follows the same logical process regardless of the place, time, and risk class.

Declarations

Conflicting interests

The authors declare no conflict of interest.

Funding

This research received no external funding.

Ethical approval

Institutional ethics approval was attained on 22/02/2018, reference id: 19925.

Guarantor

O.A.M

Contributorship

Conceptualization, J.S., V.E ., S.D., and O.A.M.; formal statistical analysis, O.A.M.; software development, O.A.M.; software logo design, O.A.M.; technical documentation preparation, O.A.M; manuscript draft writing, O.A.M; supervision, J.S., S.D., V.E.; All authors have read and agreed to the published version of the manuscript.

Acknowledgement

The authors would like to thank Mrs Lesley Hutchins, BU Innovations Limited (BUI), Professor Peter Thomas, Dr Elizabeth Williams, Ms Carla Smith, Ms Lisa Cranham, Ms Rosemary Collins, Mr Derek Kelly, Dr Juan Campos-Perez, The Research Development and Support (RDS) at Bournemouth University, and Oxentia Oxford’s Global Innovation consultancy for their advice and support.

Abbreviations

Acronyms & abbreviations	Definition
App	Application
BUI	BU Innovations Limited
CDS	Clinical decision-support
CE	Conformité Européenne
CRAN	Comprehensive R archive network
DoC	Declaration of conformity
DoS	Denial-of-service
EC	European community
EU	European union
GI	Gastrointestinal
GPL	General public license
Hb	Haemoglobin concentration
HTTP	Hypertext transfer protocol
IDA	Iron deficiency anaemia
IDIOM	Iron Deficiency as an Indicator of Malignancy
KLM	Keystroke level model
MCV	Mean cell volume
MDD	Medical device directive
MDR	Medical device regulation
MHRA	Medicines and Healthcare Products Regulatory Agency
PMS	Post-market surveillance
SAMD	Software as a medical device
SSL	Secure sockets layer
UKCA	United Kingdom conformity assessed
URL	Uniform resource locator

References

1. Hung N, et al., Risk of Cancer in Patients with Iron Deficiency Anemia: A Nationwide Population-Based Study. PLoS One, 2015. 10(3).
2. Silva, A.C., et al., Clinical risk factors for underlying gastrointestinal malignancy in iron deficiency anaemia: the IDIOM study. Frontline Gastroenterol, 2014. 5(4): p. 237-242.
3. Stone H, et al. The Dedicated Iron Deficiency Anaemia Clinic: a 15-Year Experience. 2020. DOI: https://doi.org/10.1136/flgastro-2020-101470.
4. Almilaji O, et al. Refinement and Validation of The Idiom Score for Predicting the Risk of Gastrointestinal Cancer in Iron Deficiency Anaemia. 2020. 7, DOI: https://doi.org/10.1136/bmjgast-2020-000403.
5. Wijayasekara C, et al., Clinical Risk Factors for Underlying Gastrointestinal Malignancy in Iron Deficiency Anaemia–Prospective Validation of The IDIOM Score. Cogent Medicine, 2016. 3.
6. Almilaji O, et al., Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia. Nature Scientific reports, 2021.
7. National Institute for Health and Care Excellence (NICE). Suspected Cancer: Recognition and Referral. 2015, updated Sep 2020; Available from: https://www.nice.org.uk/guidance/ng12.
8. Almilaji O, Snook J, and Thomas P, The IDIOM App (ver 1.0): A Web-based Application to Predict the Risk of Gastrointestinal Cancer in Iron Deficiency Anaemia (https://www.predict-gi-risk-in-ida.com). [Software Medical Device, Class I, MDD]. BU Innovations Limited (BUI). 2020: Bournemouth, UK.
9. Spiro, Erwin. “The Incidence of Time in the Conflict of Laws.” The International and Comparative Law Quarterly, vol. 9, no. 3, 1960, pp. 357–382. JSTOR, www.jstor.org/stable/755964.
10. Medical Device Directive (MDD) EU 93/42/EEC of 14 June 1993 concerning medical devices as amended by Directive 2007/47/EC.
11. Medical Device Regulation (MDR) EU 2017/745.

12. IEC 62304: Medical Device Software – Software Life Cycle Processes. The International Electrotechnical Commission. Published 2006.
13. IEC 62366: Medical Devices Application of usability engineering to medical devices. The International Electrotechnical Commission. Published 2005.
14. ISO 14971: Medical devices – Application of risk management to medical devices. The International Organization for Standardization. Published 2019.
15. ISO 9241-210. Ergonomics of human-system interaction: Human-centered design for interactive systems. The International Organization for Standardization. Published 2019.
16. ISO 13485. Medical devices quality management. The International Organization for Standardization. Published 2016.
17. Implementation of risk management principles and activities within a Quality Management System. The Global Harmonization Task Force (GHTF). Published 2005.
18. Clinical Risk Management: its Application in the Manufacture of Health IT Systems (DCB0129). NHS Digital Clinical Safety team. Published under section 250 of the Health and Social Care Act 2012 and released on June 2018.
19. Clinical Risk Management: its Application in the Deployment and Use of Health IT Systems (DCB0160). NHS Digital Clinical Safety team. Published under section 250 of the Health and Social Care Act 201 and released on June 2018.
20. Classification of medical devices. European Commission Health and Consumer Directorate B. Medical Devices Documents (MEDDEV 2. 4/1 Rev. 9). Published 2010.
21. Guidelines on the Qualification and Classification of Stand-Alone Software Used in Healthcare within the Regulatory Framework of Medical Devices. Medical Devices Documents (MEDDEV 2. 1/6). European Commission Directorate Consumer Environmental and Health Technologies. Published 2016.

22. Medical Device Stand-Alone Software Including Apps. The UK Medicines and Healthcare Products Regulatory Agency (MHRA). Published 2014.
23. Software as a Medical Device (SAMD): Key Definitions. International Medical Device Regulators Forum (IMDRF). Published 2013.
24. Software as a Medical Device (SAMD): Clinical Evaluation. International Medical Device Regulators Forum (IMDRF). Published 2017.
25. Principles of Labelling for Medical Devices and IVD Medical Device. International Medical Device Regulators Forum (IMDRF). Published 2019.
26. Brexit: business guidance: Regulating medical devices in the UK; What you need to do to place a medical device on the Great Britain, Northern Ireland and European Union (EU) markets. Medicines and Healthcare products Regulatory Agency (MHRA). Published 2020
27. Almilaji O, Engen V, Thomas P, et al. O67 The development of a web-based application to predict the risk of GI cancer in IDA Gut 2021;70:A37-A38. DOI: 10.1136/gutjnl-2020-bsgcampus.67.
28. Almilaji O, Thomas P, Snook J. PWE-042 Predicting the risk of gastrointestinal cancer in iron deficiency anaemia. Gut 2019;68:A192. DOI: 10.1136/gutjnl-2019-BSGAbstracts.366.
29. Almilaji O, Parry S, Thomas P, et al. P288 Downstaging of right-sided colorectal cancer diagnosed through iron deficiency anaemia. Gut 2021;70:A190. DOI: 10.1136/gutjnl-2020-bsgcampus.362.
30. Almilaji, O. Virtual workshop between Bournemouth University (UK) and Northeastern University (China), IDIOM: A Software Medical Device to Predict the risk of GI cancer in IDA. 4 Dec 2020.
31. Almilaji, O., Showcasing IDIOM App, in Transform Healthcare through innovation. April 2019: Bournemouth University, UK.
32. Almilaji, O., 2021. Modelling the episodes of care for IDA patients in a secondary-care centre using continuous-time multistate Markov chain. SJG (in press).

33. Predict Prostate Homepage. Available from: https://prostate.predict.nhs.uk/. Accessed 31 Aug 2021.
34. Filipe, J., Maciaszek L. A. Evaluation of Novel Approaches to Software Engineering. ISBN: 9783642540929, 3642540929. Springer. Dec 2013.
35. Capurso, G., Baccini, F., & Osborn, J. 2004. Can patient characteristics predict the outcome of endoscopic evaluation of iron deficiency anemia: a multiple logistic regression analysis. Gastrointestinal Endoscopy. 59:766-771.
36. Ho, C. H., Chau, W. K., Hsu, H. C. 2005. Predictive risk factors and prevalence of malignancy in patients with iron deficiency anemia in Taiwan. American Journal of Hematology. 78:108-112.
37. James, M. W., Chen, C. M., Goddard, W. P. 2005. Risk factors for gastrointestinal malignancy in patients with iron-deficiency anaemia. European Journal of Gastroenterology & Hepatology. 17:1197-1203.
38. Wijayasekara, C. M., Surgenor,S. L., Hatcher, C., Clegg, A. J., Williams, E. J., & Snook, J.2016 Clinical Risk Factors for Underlying Gastrointestinal Malignancy in Iron Deficiency Anaemia–Prospective Validation of the IDIOM Score, Cogent Medicine, 3: 1, 1213216, doi:10.1080/2331205X.2016.1213216.
39. Altman, D. G., & Royston, P. 2006. The cost of dichotomising continuous variables. BMJ (Clinical research ed.), 332(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080.
40. Almilaji, O., Webb, G., Maynard, A., Chapman, T. P., Shine, B., Ellis, A. J., Hebden, J., Docherty, S., Williams, E. J., Snook, J. 2021. Broad external validation of a multivariable risk prediction model for gastrointestinal malignancy in iron deficiency anaemia. Diagnostic and prognostic research, 5(1), 23. https://doi.org/10.1186/s41512-021-00112-8.
41. Stone, D., Jarrett, C., Woodroffe, M., Minocha, S. User Interface Design and Evaluation. ISBN: 9780080520322, 0080520324. Elsevier Science. Apr 2005.

42. Snook J, Bhala N, Beales ILP, Cannings D, Kightley C, Logan RPH, Pritchard M, Sidhu R, Surgenor S, Thomas W, Verma AM, Goddard AF. 2021. British Society of Gastroenterology guidelines for the management of iron deficiency anaemia in adults. Gut. September. doi: 10.1136/gutjnl-2021-325210N, et al., Risk of Cancer in Patients with Iron Deficiency Anemia: A Nationwide Population-Based Study. PLoS One, 2015. 10(3).

Chapter 8 : Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia

Open access original research article published at: https://www.nature.com/articles/s41598-021-92623-z

Citation: Almilaji, O., Parry, S.D., Docherty, S. et al. Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia. 2021. *Scientific Reports;* 11, 13055. https://doi.org/10.1038/s41598-021-92623-z

Abstract

Iron deficiency anaemia (IDA) is common in colorectal cancer (CRC), especially, in right-sided CRC which is known to have an overall worse prognosis. The associations between diagnostic pathway (Bowel Cancer Screening Programme (BCSP), IDA, symptomatic) and tumour side/stage was assessed using logistic regression models in 1138 CRC cases presenting during 2010-2016 at a single secondary-care centre in the UK. In the IDA sub-group, the relationship between CRC stage and the event of having a blood count prior to CRC diagnosis was examined using Bayesian parametric survival model. IDA was found as the only significant predictor of right-sided CRC (OR 10.61, 95% CI 7.02 - 16.52). Early-stage CRC was associated with both the IDA (OR 1.65, 95% CI 1.18-2.29) and BCSP pathway (OR 2.42, 95% CI 1.75-3.37). At any age, the risk of detecting CRC at late-stage was higher in those without a previous blood count check (hazard ratio 1.53, 95% credibility interval 1.08-2.14). The findings of this retrospective observational study suggest a benefit from diagnosing CRC through the detection of IDA, and warrant further research into the prognosis benefit of systematic approach to blood count monitoring of the at-risk population.

Keywords: Iron deficiency anaemia; colorectal cancer; stage; side or location; pathway or presentation.

Abbreviations

BCSP	Bowel Cancer Screening Programme
CI	Confidence Interval
CRC	Colorectal Cancer
GP	General Practitioner
Hb	Blood Haemoglobin Concentration
IDA	Iron Deficiency Anaemia
OR	Odds Ratio

Introduction

Colorectal cancer (CRC) is the fourth common cancer in the United Kingdom, accounting for 12% of all new cases; and the second common cause of cancer-related death, responsible for about 10% of all cancer deaths in the UK [1, 2]. Although the outlook is slowly improving, the 5-year survival rate for CRC is still relatively poor at 58% because most CRC cases in the UK are diagnosed at late stage [3, 4].

It has been recognized that those with more advanced CRC at diagnosis have a worse prognosis, leading to the development of the TNM staging system for CRC [5]. The association is striking – treated five-year survival ranges from over 90% for stage I disease down to about 10% for stage IV disease [1, 4]. The fact that tumour stage generally increases progressively with time, highlighting the importance of early diagnosis. Unfortunately, CRC may not cause symptoms until the disease is already advanced, and when symptoms do develop there is sometimes reluctance to seek medical advice. The consequence of these delays is that many cases of CRC present at a late stage, with a correspondingly high mortality rate. The focus over recent years has therefore been on early diagnosis by screening of the pre-symptomatic at-risk population [1, 6].

The English Bowel Cancer Screening Programme (BCSP) was developed with the aim of reducing the mortality rate by both earlier detection of CRC and removing polyps which if left untreated might advance to cancer [6]. The BCSP is based on the biennial offer of a faecal occult blood test to all in the population aged 60 – 74, with a view to colonoscopy if positive.

Bowel cancer screening has been shown to reduce the mortality rate of CRC by about 15% with faecal occult blood testing [7, 8], probably because cases were detected at an earlier stage [1, 9]. The proportion of CRCs diagnosed at

early stage (I or II) was about 64% for the BCSP in 2017, compared to 47% for GP referrals and 32% for those presenting as emergency admissions [10]. However only about 10% of all CRCs countrywide are detected through the BCSP [11]. The relatively low proportion of screened detected cancers probably relates to a number of factors, including low population uptake (less than 50% in some areas) and limited sensitivity of the initial screening test [9].

Overall about a third of CRCs occur in the right colon, and these differ in a number of important respects from those found in the left colon [1, 12]. Right-sided CRCs tend to present with larger tumours at a more advanced stage, and a correspondingly worse prognosis [12-18]. They are also strongly associated with the finding of iron deficiency anaemia (IDA) at presentation [19-23], believed to be due to chronic low-grade loss of (iron-rich) blood from the tumour bed, resulting in the slowly progressive depletion of body iron stores. IDA often occurs before any other clinical manifestations of CRC [24], and as the development of IDA is gradual it may precede the diagnosis of CRC by up to 2 years [25]. This provides a window of opportunity for the detection of CRC earlier in the disease course, particularly for tumours of the right colon, and is the basis of the recommendation for urgent investigation of unexplained IDA in the at-risk population [26, 27].

The study reported here is based on the analysis of a large dataset of patients with CRC presented through different diagnostic pathway at a single centre, and the objectives were twofold. First, to compare the effect of the three major diagnostic routes for CRC – the IDA, BCSP, and symptomatic pathways – on the stage and side of CRC. Second, to explore the scale of the missed opportunity for earlier diagnosis of CRC through the IDA pathway, by assessing the prevalence and results of blood counts prior to CRC diagnosis in the IDA pathway sub-group, and comparing the relationship between prior blood count event and the risk of late-stage disease at diagnosis.

Methods

This study is a retrospective observational study involved statistical analysis of anonymised secondary clinical data on the Poole Hospital CRC MDT database for the years 2010 to 2016 inclusive. Assuming the smallest effect size (0.1), and significance level= 0.05, the sample size was estimated to be around 967 when power= 80%, and around 1268 when power= 90%. The *Strengthening the Reporting of Observational Studies in Epidemiology* (STROBE) guidelines were used to ensure the reporting of this study. Since this is an observational study, and simply involved the analysis of anonymised secondary clinical data, formal ethical/institutional approvals, consent to participate/publish were not required.

The association between stage/ side and presentation pathway

The first part of this study involved the statistical analysis of 1258 CRC cases. The data was scrutinized in 2018 for the purposes of a service audit and included:

- age at diagnosis
- sex
- haemoglobin concentration (Hb) at presentation
- presentation pathway (IDA, BCSP or symptomatic)
- tumour stage (of the most advanced if synchronous lesions present)
- tumour number, histology and location(s)

Iron deficiency was defined by transferrin saturation <15% and / or serum ferritin less than the lower laboratory limit of normal at the time of the analysis. The symptomatic group comprised cases with symptoms relating directly to the underlying CRC (other than symptomatic anaemia) that resulted in GP referral or emergency admission to secondary care. Patients with both bowel symptoms and IDA were allocated to a presentation pathway based on which

was felt to be the dominant feature – in a few cases this was rather arbitrary, but the allocation was made without knowledge of tumour site or stage.

The diagnosis of CRC was established by standard clinical investigation including colonoscopy, and CT scanning. Tumours were graded according to the simplified TNM staging system [5] based on the initial radiological appearances, modified in the light of subsequent surgical and pathological findings where available. For the purposes of the analysis, stage I and II CRCs were combined into one category - early stage; whilst stage III and IV CRCs were combined as a second category - late stage. CRCs located at or beyond the splenic flexure were considered left-sided, and those proximal to splenic flexure right-sided. Eight cases had synchronous CRCs, and for the purposes of this study they were considered right-sided if any tumour was proximal to the splenic flexure.

The exclusion criteria were (a) incomplete records (17 cases), (b) second entry due to metachronous CRC (7), (c) other neoplastic diagnoses such as stromal tumours, small bowel carcinoma, neuroendocrine tumours, and anal carcinoma (35), (d) non-incident presentation / diagnosis made at another hospital (27), and (e) diagnosis of CRC on cancer follow-up or as an incidental finding on a scan undertaken for some unrelated reason (34). When no histological confirmation was found, cases were included only if the radiological features were regarded as characteristic of CRC, and they were managed as such clinically.

The effects of age, sex, Hb, and presentation pathway on tumour stage (early / late) or side (left / right) were analysed using simple binary logistic regression models run for each of the predictors separately, with stage or side as the outcome. When any significant association was found ($p < 0.05$), the predictor was added to a multivariable logistic regression model. Due to correlations with particular presentation pathways (such as in the case of BCSP and age, and IDA and Hb), only simple regression models were built for age and Hb.

Statistical methods used to check the validity of the fitted logistic regression models and the goodness of fit are shown in (Table S 8-1, supplementary information).

The association between prior blood test event and stage

The second part of the study involved a detailed assessment of the 171 IDA sub-group from all the 1258 cases dataset. An arbitrary “presentation period” was defined as the three months immediately prior to the date of CRC diagnosis. The anonymised data for each subject included whether a blood count had been checked in the 3 years prior to the start of their presentation period, and if so, the date and Hb result for the last blood count in this window. On the basis of published literature regarding temporal changes in blood count prior to the diagnosis of CRC [26], an arbitrary window of 2 years was taken as the basis of comparison for this study.

A proportional hazards parametric survival model was employed to estimate the effect of previous blood count testing (done / not done) on the onset time of late-stage disease in the IDA sub-group using current status data. Current status data consisted of (a) observation time (CRC diagnosis time) and (b) whether the observation time was smaller or larger than the time to late-stage CRC. Diagnosis time was assumed to be independent of late-stage CRC onset time, and survival time (free of late-stage CRC ie diagnosed with early stage CRC) to equal age (in years).

The endpoint of interest was “time to late-stage CRC”. So, if patient ***i*** was investigated at age Ci and late-stage CRC diagnosed, the time of onset was recorded as the interval [0, Ci]. If early-stage CRC was diagnosed, then the time of late-stage onset was recorded as the interval [Ci, ∞].

The Weibull distribution was specified as the baseline parametric distribution because it allows for constant, increasing, or decreasing hazard rates. To approximate the posterior distribution parameters, four Markov Chain Monte Carlo (MCMC) methods were used (sample size per chain was 1000). As current status data was uninformative, we incorporated prior information into the analysis by extending the parametric model to Bayesian framework. The prior information was based on the following assumptions (a) hazard rates of late-stage disease do not decrease with age and (b) without intervention, all early-stage CRCs would progress to late-stage within 10 years. To incorporate (a) the shape parameter was constrained to be >1, whilst for (b), we set a maximum possible time equal to the age of the patient plus 10 years instead of the upper end of the interval (∞). Statistical assessments of validity and goodness of fit of the models were based on the method outlined in (Table S 8-1, Figure S 8-1, Figure S 8-2, Supplementary information).

The methods used in this study were guided by previous relevant publications [28-32]. R (version 3.6.1) and RStudio (version 1.2.5001) were used to run the statistical analyses and to produce the descriptive statistics, and graphs.

Ethics declarations. Retrospective analysis of anonymised secondary data, formal research ethics approval was not required.

Results

The association between stage/ side and presentation pathway

After tidying the database and applying the exclusion criteria, 1138 complete cases were available for detailed analysis. Of these, 90% had histologically confirmed colorectal adenocarcinoma, and most of the remainder had high-grade dysplasia on biopsy, undifferentiated carcinoma, or signet cell carcinoma. As shown in Table 8-1, almost 70% of cases presented via the symptomatic pathway, with about 15% each through the IDA and BCSP

routes. Overall, 45% of cases presented with early-stage disease, and 39% with right-sided tumours.

Table 8-1 Descriptive statistics of the CRC dataset divided by presentation pathway (IDA – iron deficiency anaemia, BCSP – Bowel Cancer Screening Programme, and symptomatic group)

		IDA	BCSP	Symptomatic
Number	n (%)	171 (15.0%)	187 (16.4%)	780 (68.6%)
Sex ratio	M / F	1.1	1.5	1.3
Age (years)	Median (Q1 - Q3)	78 (71 - 86)	68 (64 - 71)	75 (64 - 83)
Hb at diagnosis (g/l)	Median (Q1 - Q3)	89 (80 - 100)	138 (126 - 147)	124 (106 - 140)
Early stage (I +II)	n (%)	89 (52.0%)	115 (61.5%)	310 (39.7%)
Right-sided	n (%)	141 (82.5%)	56 (29.9%)	245 (31.4%)

As anticipated, the BCSP group were more likely to be male, and to be younger. The proportion with right-sided tumours was markedly higher in the IDA group and slightly reduced in the BCSP group.

By crude comparison with the symptomatic group, there was a greater percentage of early-stage CRCs in both of the other groups (Figure 8-1).

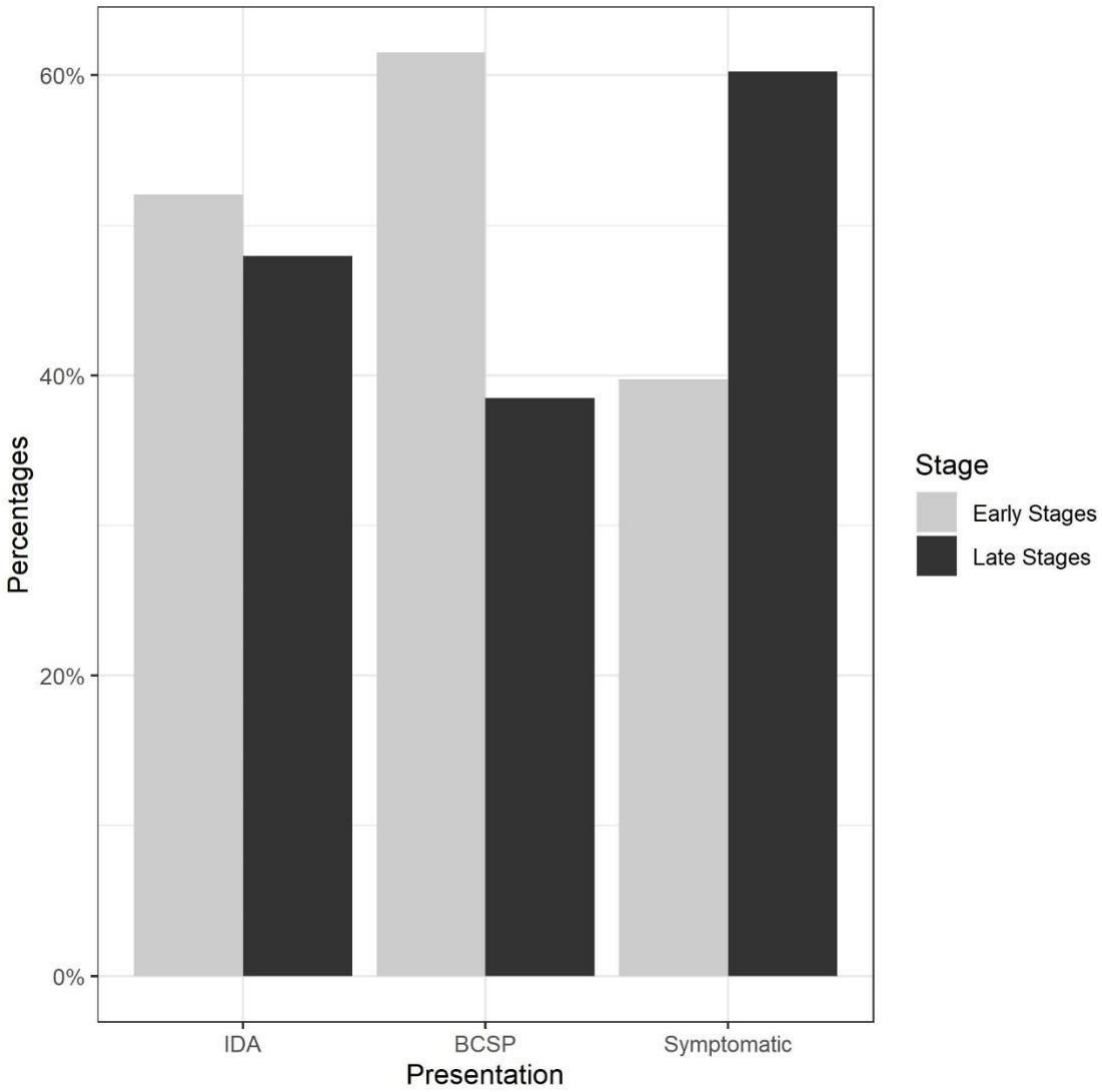


Figure 8-1 The distribution of tumour stage by presentation pathway

Four binary logistic regression models were constructed, and their findings are summarised in Table 8-2.

Table 8-2 Summary of logistic regression analyses showing variables predictive of right-sided CRC (models A – C) and early-stage CRC (model D)

Model	Outcome	Predictor	OR (95% CI)	P value
A	Right-sided CRC	Presentation (IDA)	10.61 (7.02 - 16.52)	<0.0001
		Presentation (BCSP)	0.95 (0.67 - 1.35)	0.78
		Sex (female)	1.94 (1.49 - 2.53)	<0.0001
B	Right-sided CRC	Hb (g/l)	0.95 (0.94 - 0.96)	<0.0001
C	Right-sided CRC	Age (years)	1.04 (1.03 - 1.05)	<0.0001
D	Early-stage CRC	Presentation (IDA)	1.65 (1.18 - 2.29)	0.003
		Presentation (BCSP)	2.42 (1.75 - 3.37)	<0.0001

In **model A**, analysis revealed that sex and presentation pathway were both strongly significant predictors of tumour side. The final multiple binary logistic regression model was therefore constructed according to the formula (left-side CRC as reference category):

$$\log\left\{\frac{\mathbb{P}(Side = right)}{\mathbb{P}(Side = left)}\right\} = \beta_0 + \beta_1 presentation + \beta_2 sex$$

The odds of right-sided CRC were about 11 times higher for the IDA pathway than the symptomatic one, whilst the BCSP route was not a significant predictor of right-sided CRC. CRCs were 94% more likely to be right-sided in females compared to males.

In **model B**, Hb was found to be a very significant negative predictor of right-sided CRC - for each unit (g/l) decrease in Hb, there was about a 5% increase in the odds of right-sided CRC.

Model C showed that age is also a very significant positive predictor of right-sided CRC - for each rising year of age, the odds of right-sided CRC increased by about 4%.

In **model D**, statistical analysis showed that only presentation pathway was a significant predictor of early-stage CRC. The association between tumour side and stage is not statistically significant (p = 0.07). The final binary logistic regression model was therefore constructed according to the formula (late stage as reference category):

$$\log\left\{\frac{\mathbb{P}(Stage = early)}{\mathbb{P}(Stage = late)}\right\} = \beta_0 + \beta_1 presentation$$

The findings indicate that IDA was a significant positive predictor of early stage CRC. Results also show CRCs presenting through the IDA and BCSP routes are 65% and 142% respectively more likely to be diagnosed at early stage, as compared to the symptomatic pathway. Statistical assessment of validity and goodness of fit of the logistic regression models was satisfactory (based on the criteria outlined in Table S 8-1 – Supplementary information).

The association between prior blood test event and stage

Figure 8-2 shows the cumulative percentage prevalence of blood counts over the three years prior to the presentation period for CRC, for the 171 cases presenting via the IDA pathway.

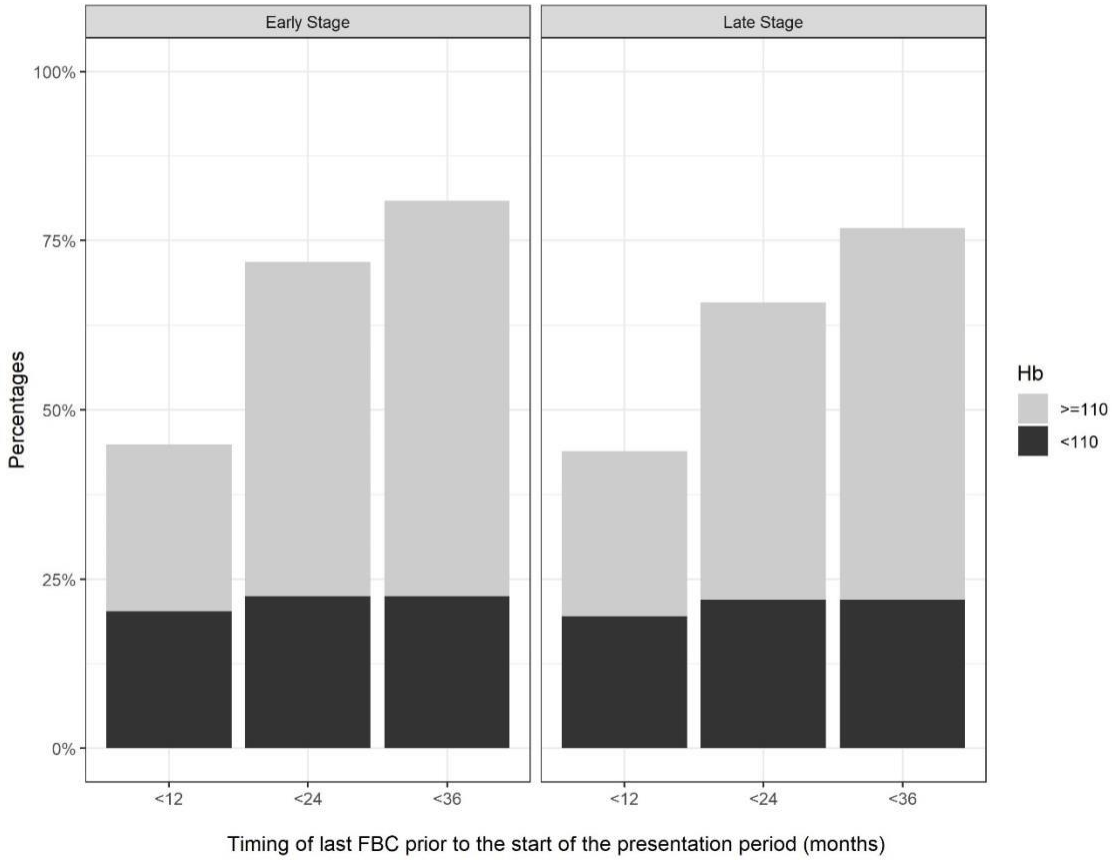


Figure 8-2 The cumulative percentage prevalence of blood count checks in the 12, 24 and 36 months prior to the diagnosis of CRC in the IDA group, sub-divided according to the tumour stage (early / late) and Hb result (g/l)

In the two years prior to diagnosis of CRC, 31% of did not have a record of any blood count, and a further 22% had an abnormally low blood count (Hb <110g/l) which did not result in immediate referral. Most of these abnormal results were recorded in the 12 months prior to the presentation period.

Descriptive statistics for the IDA group broken down by the result of the last blood count in the two years prior to the presentation window of CRC are shown in Table 8-3. There were trends towards those with 'blood test not done' being younger, more likely to have right-sided CRC, and less likely to have early-stage disease.

Table 8-3 Descriptive statistics for the IDA group (n = 171) by outcome of the last blood count in the two years prior to presentation with CRC

		Hb > 110g/l	Hb <110g/l	Not done
Number	n (%)	80 (47%)	38 (22%)	53 (31%)
Sex ratio	M / F	1.5	0.7	1.1
Age (years)	Median (Q1 - Q3)	78 (75 - 85)	83 (77 - 88)	73 (66 - 83)
Hb at diagnosis (g/l)	Median (Q1 - Q3)	94 (85 - 102)	84 (71 - 92)	87 (74 - 99)
Early stage	n (%)	44 (55.0%)	20 (52.6%)	25 (47.2%)
Right-sided	n (%)	65 (81.2%)	30 (78.9%)	46 (86.8%)

Bayesian Weibull regression showed that the posterior baseline survival distribution of IDA patients with early-stage CRC at diagnosis (ie not having reached late-stage disease) decreased with increasing age. This survival figure fell from 80% at age 60 to about 35% at age 80 (Figure 8-3a); a. Analysis also revealed that having a prior blood test (regardless of result) was significantly related to time to late-stage disease (mean (sd): 0.66 (0.18), 95% credibility interval: 0.46-0.93). Looked at the other way, the hazard ratio for detecting CRC at late-stage was 53% higher in those without a previous blood count (1.53, 95% credibility interval 1.08 – 2.14).

The posterior median onset time of late-stage CRC in those with a blood count in the preceding two years was 75 years (95% credibility interval: 72 – 78). This was 5 years later than the median onset time of 70 years of age (95% credibility interval: 65 – 74) for those without a blood count. This implies that

at a given age, the probability that CRC is detected at an early-stage is higher in those with a previous blood count (Figure 8-3b).

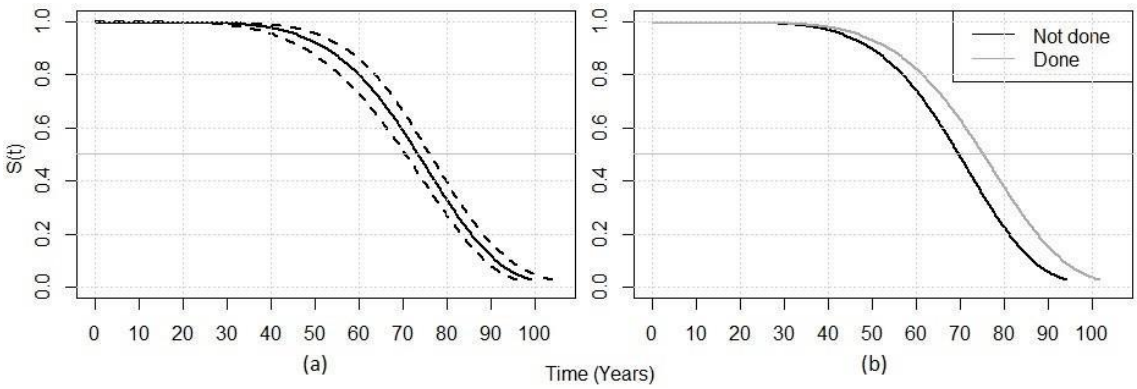


Figure 8-3 Posterior survival distributions for the IDA group, showing (a) the baseline survival probability at any given time S(t), with dashed lines representing the credibility interval, and (b) the survival probabilities for sub-groups categorised by whether a blood count was done in the two years prior to the presentation window

Discussion

Our results demonstrate that during the seven-year study period, just over 30% of CRCs were diagnosed via either the IDA or BCSP pathway, with similar numbers in each. Comparison with national data reveals a similar proportion of early-stage disease diagnosed through the symptomatic pathway, at around 40% [10]. The figure for the BCSP pathway is slightly lower than the national figure (62% v 64%), but this may reflect the higher proportion of right-sided cancers detected (30% v 23%) [9, 10].

We have confirmed previous observations that IDA is strongly associated with right-sided CRC, but the striking finding from our study is that diagnosis through a contemporary IDA pathway has the potential to downstage the disease, as previously demonstrated for the BCSP pathway [9]. This is in

contrast to reports in the literature suggesting that IDA is a marker of poor prognosis in CRC [22, 23, 33, 34]. The reasons why our findings differ from those of some historical studies may include issues of confounding and diagnostic delay. Firstly, the risk of confounding arises from the strong association between IDA and right-sided CRC – a pattern of disease which is associated with later diagnosis at a more advanced stage, and a correspondingly poorer prognosis [12-18]. Secondly, the diagnosis of CRC through the IDA pathway has been beset by delays resulting in late diagnosis of CRC and so a poor prognosis [35-39].

In years past major delays at three points in the pathway of CRC diagnosis through the detection of IDA were all too common, and the cumulative effect of these delays may have been a major contributor to the historical association of IDA with poor prognosis in CRC [35-39]. The first is confirmation of IDA on a blood test – a particular issue as even severe anaemia may not cause appreciable symptoms. The second is lack of awareness of the significance of IDA as a marker of underlying malignancy, and therefore of the importance of swift referral for investigation. The third is the time between referral and an adequate diagnostic examination of the (right) colon. Survival in anaemic CRC appears to be inversely related to this last delay [39].

Various developments over recent years have had a major bearing on these delays. Firstly, routine blood count checks in the at-risk population have become much more frequent and widespread - the rate of blood count testing increased progressively in the UK between 2000 and 2015, from approximately 160 to 430 per 1000 population per annum [40].

Secondly, much has been done to accelerate the referral and comprehensive investigation of patients found to have IDA, particularly those at risk of CRC. This includes education in primary care, national guidelines encouraging fast-track referral [26, 27], and the development of dedicated IDA triage services

in secondary care – such as the IDA Clinic at Poole [41, 42], which was incidentally operational throughout the years of this study. Finally, gastroenterology speciality groups have introduced quality initiatives to improve the diagnostic yield of investigation, particularly in the right colon [43].

The strengths of this study are the novelty of examining the association between the event of having prior blood count check and the CRC stage in IDA patients, and inclusion of a BCSP CRC group as a positive control. Limitations include the uncertain applicability of a single centre experience to other populations, and being a retrospective analysis, our inability to control the size of the study subgroups or to incorporate other variables that could impact the prognosis in CRC. In fact, the major potential constraint of the study was the use of stage/side as the only markers of prognosis in CRC. We feel however that this methodology is justified because the link between stage/side and prognosis is so strong [1, 4], and this view is supported by the results for the BCSP group, which fit well with the established improvement in prognosis with this programme [7-9, 44]. Nonetheless, further studies are clearly warranted to corroborate the findings.

With correction for confounding and a reduction in diagnostic delays, our results suggest that CRC in the right colon may be detected at earlier stage with a correspondingly better prognosis. This observation strengthens the case for the inclusion of monitoring for IDA in the repertoire of screening approaches for the early diagnosis of CRC. Currently however there is no systematic process for routinely checking blood counts in the at-risk population, despite the universal presence of the necessary laboratory infra-structure.

Bearing in mind that the development of IDA is a gradual process prior to the diagnosis of CRC, our results suggest that there may be scope for further improvement in how we screen for bowel cancer. Of those diagnosed with CRC via the IDA pathway, some 31% had not had a blood count in the two

years prior to diagnosis, whilst a further 22% had a low blood count – that in retrospect may perhaps have been indicative of undiagnosed CRC. Our results also suggest that the median onset age of late-stage CRC in those with a previous blood count is about five years older than in those without, so that for a given age, the proportion of CRCs detected at early stage is higher. This is an interesting observation for which there are various possible explanations, but we feel that confounding is perhaps the most likely - individuals who avoid medical care are inherently less likely to have a blood test, and also less likely to present early with their undiagnosed CRC.

Nevertheless, we feel that a strong case can be made for formally recommending a blood count test on perhaps an annual basis in the at-risk population – with follow-up iron studies for those with detected anaemia. Blood count checks have an advantage over the current CRC screening modalities of stool testing or sigmoidoscopy in being more acceptable to many people. However, the recommendation would be to introduce blood count checks as a screening test complementary to the current BCSP, not an alternative. The logic to this is that IDA screening would be expected to predominantly detect right-sided CRC, whilst the current BCSP predominantly targets left-sided CRC, with the suspicion that it may be less effective at picking up right-sided lesions [9, 44].

In conclusion, our findings suggest prognostic benefit from diagnosing CRC through the detection of IDA, and that IDA screening is currently sub-optimal. These observations strengthen the case for a systematic approach to blood count monitoring of the at-risk population.

Declarations

Competing interests

The authors declare that they have no conflict of interest.

Funding

PhD studentship (OAM) jointly funded by Poole Hospital Gastroenterology Research fund and Bournemouth University.

Availability of data and code

Data and code available on reasonable request.

Author' contributions

OAM, SDP, and JS conceived and designed this study. OAM analysed the data, wrote the draft, and prepared the tables and figures. All authors reviewed and made significant contributions to the subsequent revision of the paper.

Acknowledgement

The authors would like to thank Professor Peter Thomas for his contribution in this study.

References

1. Dekker E, Tanis PJ, Vleugels JLA, et al. Colorectal Cancer. Lancet. 2019; 19:1467-80.
2. Cancer Research UK. Bowel Cancer Incidence Statistics, 2017. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Zero. Accessed Oct 2, 2020.
3. Cancer Research UK. Bowel Cancer Survival Statistics, 2017. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Two. Accessed Oct 2, 2020.
4. Cancer Research UK. Bowel Cancer Survival by Stage at Diagnosis, 2018. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival#heading-Three. Accessed Oct 2, 2020.
5. Compton CC, Greene FL. The Staging of Colorectal Cancer: 2004 and Beyond. Cancer Journal for Clinicians. 2004; 54:295-308.
6. Public Health England. Bowel Cancer Screening Programme (BCSP): Overview, 2019. Available at: https://www.gov.uk/guidance/bowel-cancer-screening-programme-overview. Accessed Nov 8, 2020.
7. Towler B, Irwig L, Glasziou P, et al. A Systematic Review of The Effects of Screening for Colorectal Cancer Using the Faecal Occult Blood Test, Hemoccult. BMJ. 1998; 317:559-65.
8. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised Controlled Trial of Faecal-Occult-Blood Screening for Colorectal Cancer. Lancet. 1996; 348:1472-77.
9. Logan RF, Patnick J, Nickerson C, et al. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. Gut. 2012; 61:1439-46.
10. National Cancer Registration and Analysis Service. Routes to Diagnosis of Cancer by Stage, 2016. Available at:

http://www.ncin.org.uk/publications/routes_to_diagnosis. Accessed Oct 2, 2020.

11. Braun M, Hill J, Kuryba A, et al. National Bowel Cancer Audit, ACPGBI, 2016. Available at: https://www.acpgbi.org.uk/content/uploads/2016/07/nati-clin-audi-bowe-canc-2016-rep-v2.pdf. Accessed Nov 8, 2020.
12. Nawa T, Kato J, Kawamoto H, et al. Differences Between Right and Left-Sided Colon Cancer in Patient Characteristics, Cancer Morphology and Histology. Journal of Gastroenterology and Hepatology. 2008; 23:418-23.
13. Hansen I, Jess P. Possible Better Long-Term Survival in Left Versus Right-Sided Colon Cancer – A Systematic Review. Danish Medical Journal. 2012; 59: A4444.
14. Snaebjornsson P, Jonasson L, Jonsson L, et al. Colon Cancer in Iceland: A Nationwide Comparative Study on Various Pathology Parameters with Respect to Right and Left Tumor Location and Patients Age. International Journal of Cancer. 2010; 127:2645-53.
15. Petrelli F, Tomasello G, Borgonovo K, et al. Prognostic Survival Associated with Left-Sided Vs Right-Sided Colon Cancer: A Systematic Review and Meta-Analysis. JAMA Oncology. 2017; 3:211–19.
16. Yahagi M, Okabayashi K, Hasegawa H, Tsuruta M, Kitagawa Y. The Worse Prognosis of Right-Sided Compared with Left-Sided Colon Cancers: a Systematic Review and Meta-analysis. *J Gastrointest Surg*. 2016 Mar;20(3):648-55. doi: 10.1007/s11605-015-3026-6. Epub 2015 Nov 16. PMID: 26573851
17. Mik M, Berut M, Dziki L, et al. Right and Left-Sided Colon Cancer Clinical and Pathological Differences of The Disease Entity in One Organ. Archives of Medical Science. 2017; 13:157–162.
18. Ghazi S, Lindforss U, Lindberg G, et al. Analysis of Colorectal Cancer Morphology in Relation to Sex, Age, Location, And Family History. Journal of Gastroenterology. 2012; 47:619-34.

19. Beale A, Penney M, Allison M. The Prevalence of Iron Deficiency Among Patients Presenting with Colorectal Cancer. Colorectal Disease. 2005; 7:398-402.
20. Edna TH, Karlsen V, Jullumstrø E, et al. Prevalence of Anaemia at Diagnosis of Colorectal Cancer: Assessment of Associated Risk Factors. Hepatogastroenterology. 2012; 59:713-16.
21. Ho CH, Yu YB, Wu PH. The Prevalence of Iron Deficiency Anemia And Its Clinical Implications in Patients with Colorectal Carcinoma. Journal of the Chinese Medical Association. 2008; 71:119-22.
22. Alexiusdottir KK, Möller PH, Snaebjornsson P, et al. Association of Symptoms of Colon Cancer Patients with Tumor Location and TNM Tumor Stage. Scandinavian Journal of Gastroenterology. 2012; 47:795-01.
23. Tokunaga R, Nakagawa S, Miyamoto Y, et al. The Impact of Preoperative Anaemia and Anaemic Subtype on Patient Outcome in Colorectal Cancer. Colorectal Disease. 2019; 21:100-109.
24. 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 Gastroenterol. 2020; 7: e000403. Available at: https://doi.org/10.1136/bmjgast-2020-000403.
25. Edgren G, Bagnardi V, Bellocco R, et al. Pattern of Declining Hemoglobin Concentration Before Cancer Diagnosis. International Journal of Cancer. 2010; 127:1429-36.
26. Goddard A, James M, McIntyre A, et al. Guidelines for The Management of Iron Deficiency Anaemia. Gut. 2011; 60:1309-1316.
27. National Institute for Health and Care Excellence (NICE). Suspected Cancer: Recognition and Referral, updated Sep 2020. Available at: https://www.nice.org.uk/guidance/ng12. Accessed Nov 8, 2020.
28. Agresti A. Categorical Data Analysis. New Jersey: John Wiley & Sons Inc; 2002.
29. Hoel DG, Walburg HE. Statistical Analysis of Survival Experiments. Journal of the National Cancer Institute. 1972; 49:361–72. https://doi.org/10.1093/jnci/49.2.361.

30. Anderson-Bergman C. Bayesian Regression Models for Interval-censored Data in R. The R Journal. 2017; 9:487-98.
31. Anderson-Bergman C. Using Icenreg For Interval Censored Data in R V2.0.9, 2020. Available at: https://cran.r-project.org/web/packages/icenReg/vignettes/icenReg.pdf. Accessed Nov 8, 2020.
32. Moore FM. Applied Survival Analysis Using R. ISBN13 9783319312439. Springer International Publishing AG. Switzerland; 2016.
33. Stapley S, Sharp D, Hamilton W. The Mortality of Colorectal Cancer in Relation to The Initial Symptom and To the Duration of Symptoms: A Cohort Study in Primary Care. British Journal of Cancer. 2006; 95:1321–25.
34. Wilson MJ, van Haaren M, Harlaar JJ, et al. Long-Term Prognostic Value of Preoperative Anemia In Patients with Colorectal Cancer: A Systematic Review and Meta-Analysis. Surgery Oncology. 2017; 26:96-104.
35. Stebbing JF, Nash AG. Avoidable Delay in The Management of Carcinoma of The Right Colon. Annals of the Royal College of Surgeons of England. 1995; 77: 21-23.
36. Goodman D, Irvin TT. Delay in The Diagnosis and Prognosis of Carcinoma of The Right Colon. British Journal of Surgery. 1993; 80:1327-1329.
37. Acher P, Al-Mishlab T, Rahman M, et al. Iron-Deficiency Anaemia and Delay in The Diagnosis of Colorectal Cancer. Colorectal Disease .2003; 5:145-8.
38. Harris G, Simson J. Causes of Late Diagnosis in Cases of Colorectal Cancer Seen in A District General Hospital Over A 2-Year Period. Annals of the Royal College of Surgeons of England. 1998; 80:246–8.
39. Teng C-LJ, Yu J-T, Chen Y-H, et al. Early Colonoscopy Confers Survival Benefits on Colon Cancer Patients with Pre-Existing Iron Deficiency Anemia: A Nationwide Population-Based Study. PLoS

ONE. 2014; 9(1): e86714. Available at: https://doi.org/10.1371/journal.pone.0086714.

40. O'Sullivan JW, Stevens S, Hobbs FDR, Salisbury C, Little P, Goldacre B, Bankhead C, Aronson JK, Perera R, Heneghan C. Temporal trends in use of tests in UK primary care, 2000-15: retrospective analysis of 250 million tests. *BMJ*. 2018 Nov 28;363:k4666. doi: 10.1136/bmj.k4666. Erratum in: BMJ. 2019 Jan 29;364:l444. PMID: 30487169; PMCID: PMC6260131.
41. Surgenor S, Kirkham S, Parry S, et al. The Development of a Nurse-Led Iron Deficiency Anaemia Service in A District General Hospital. Frontline Gastroenterol. 2014; 5(3):219-223. https://doi.org/10.1136/flgastro-2013-100385.
42. Stone H, Almilaji O, John C, et al. The Dedicated Iron Deficiency Anaemia Clinic: a 15-Year Experience. Frontline Gastroenterology. 2020. Available at: https://doi.org/10.1136/flgastro-2020-101470.
43. Rees CJ, Gibson ST, Rutter MD, et al. UK Key Performance Indicators & Quality Assurance Standards for Colonoscopy, 2019. BSG Available at: https://www.bsg.org.uk/wp-content/uploads/2019/12/UK-Key-Performance-Indicators-and-Quality-Assurance-Standards-for-Colonoscopy-1.pdf. Accessed Nov 8, 2020.
44. Atkin WS, Edwards R, Kralj-Hans I, et al. UK Flexible Sigmoidoscopy Trial Investigators. Once-Only Flexible Sigmoidoscopy Screening in Prevention of Colorectal Cancer: A Multicentre Randomised Controlled Trial. Lancet. 2020; 375:1624-33.

Supplementary information

Table S 8-1 Statistical assessment methods employed

Analysis	Methods
The association between stage/ side and presentation pathway	Smoothed scatter plot, Cook's distance and standardised residual errors, variance inflation factor, Akaike information criterion, analysis of variance, $\chi 2$ test and pseudo R2 were used to check the validity of the fitted logistic regression models and the goodness of fit.
The association between prior blood test event and stage	To inspect the Bayesian Weibull model fit, we examined the baseline distribution by plotting it against the semi-parametric estimate to see whether there were systematic deviations for the chosen parametric distribution from the semi-parametric (Figure S 8-1). Trace plots were used to confirm that the model converged to the target distribution (Figure S 8-2), and that the sampled values for each parameter in the chain were overlapping with values close to 1.

Figure S 8-1 Demonstration of the baseline distribution against the non-parametric estimate and other parametric families

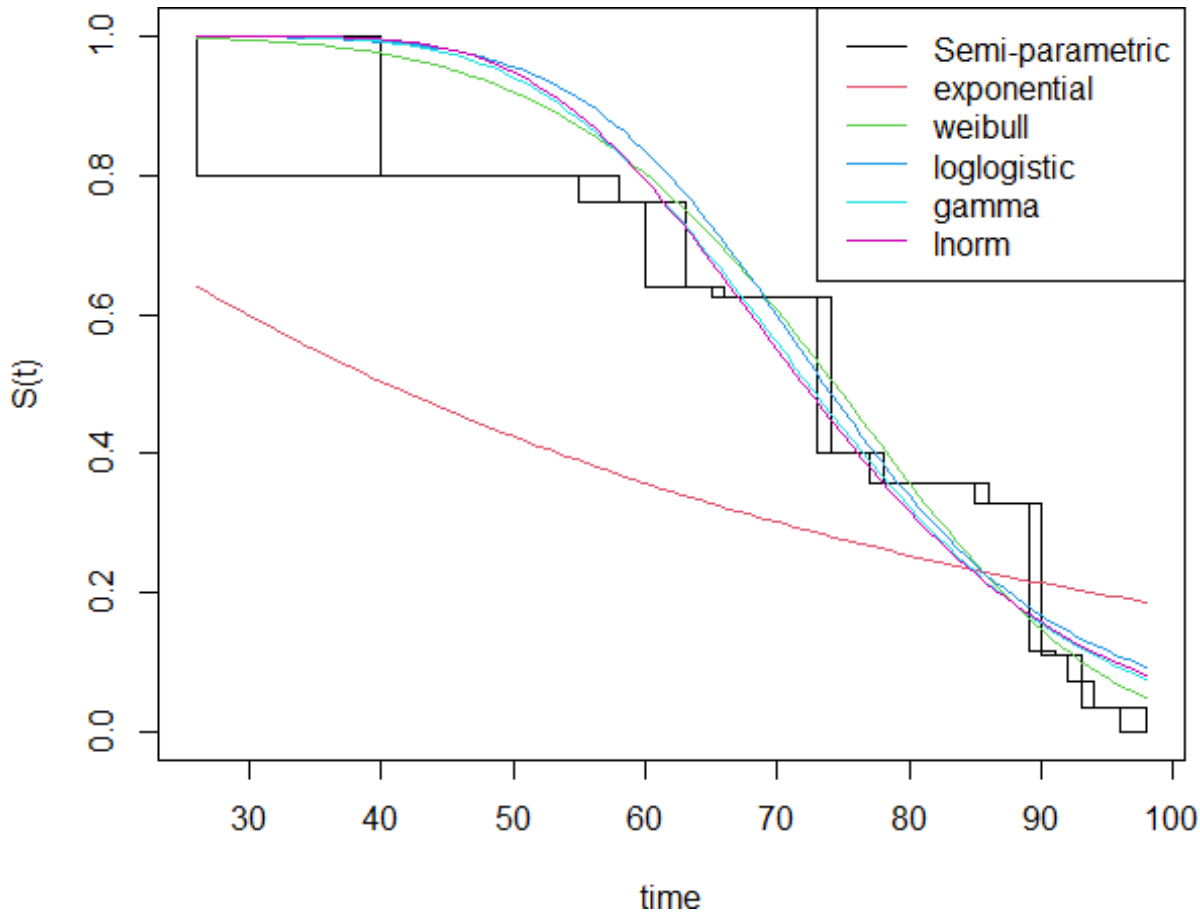
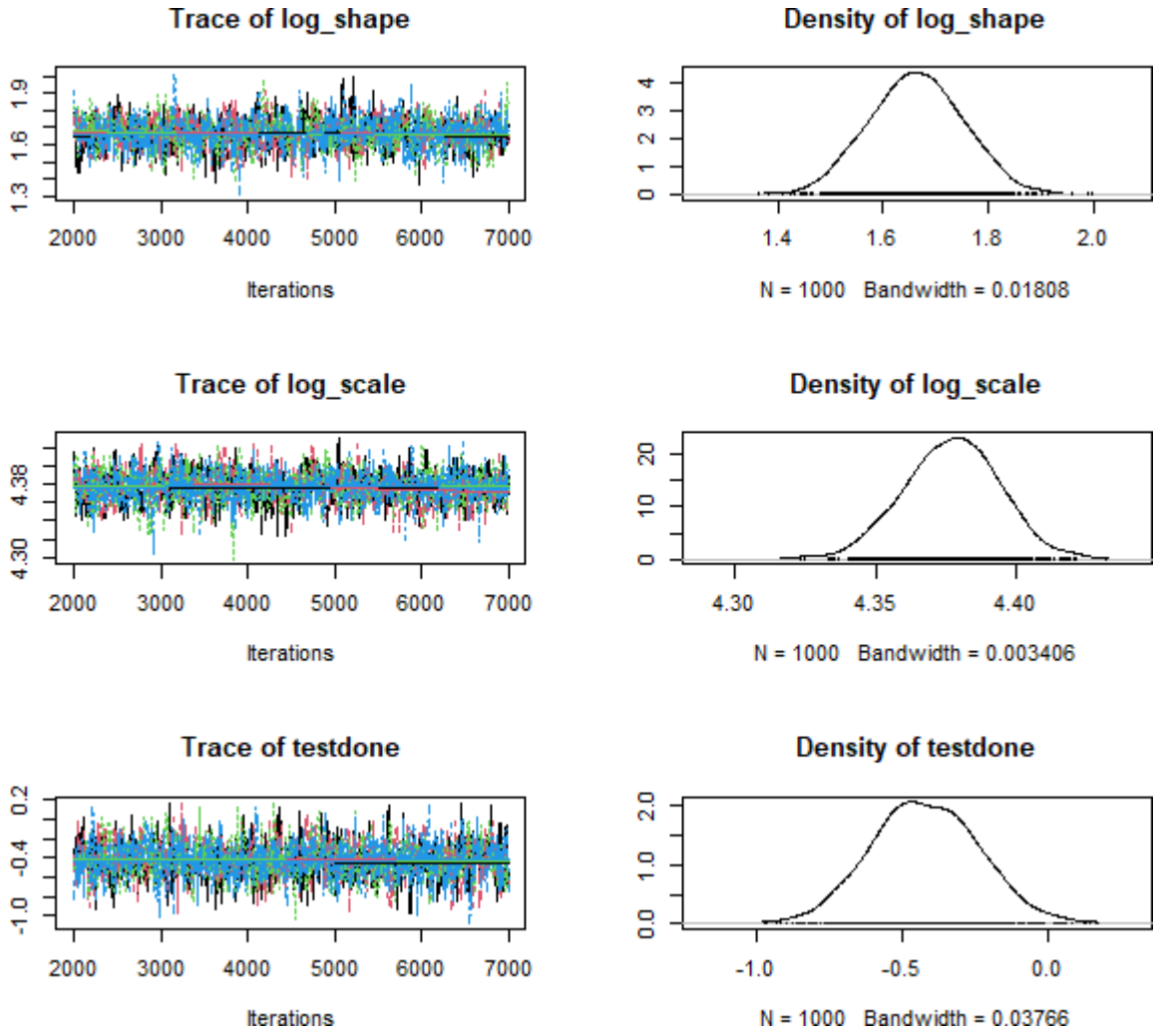


Figure S 8-2 Demonstration of the trace plots and marginal density estimates of the posterior samples



Chapter 9 : Colorectal cancer and the blood loss paradox

Original research article published as non-open access at: https://fg.bmj.com/content/early/2021/10/20/flgastro-2021-101959

Citation: Almilaji O, Parry SD, Docherty S, et al. Colorectal cancer and the blood loss paradox. 2021. *Frontline Gastroenterology*. doi: 10.1136/flgastro-2021-101959.

A pre-publication version of this chapter is available on BURO: https://eprints.bournemouth.ac.uk/36213/3/Colorectal%20cancer%20and%20the%20blood%20loss%20paradox.pdf

Chapter 10 : General discussion

10.1 Outline

This chapter starts by discussing the contribution to knowledge and key findings of this PhD study with reference to the study' objectives -set in Chapter 1-, while considering the limitations and results from similar studies. The chapter then moves into reviewing the significance of the research findings to the clinical practice. After this, it highlights its academic and societal impacts. Finally, it concludes with a summary of the PhD project and potential future research plans.

10.2 Key findings

As mentioned previously in chapter one, the analysis of this thesis starts in chapter 4, hence, the key results of this PhD project will be reviewed in the following sections, commencing with chapter 4:

Chapter 4. Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia.

This study (Almilaji et al. 2020) examined and internally validate the association between GI cancer and IDA; and showed that FIT is a predictor of GI cancer in the IDA.

By utilizing a large dataset, and by keeping the continuous exploratory variables (age, Hb, and MCV) without categorisation in the model, MCV was found as a new significant predictor of GI cancer risk in IDA patients (objective 1). The odds ratio for MCV was similar to that of Hb, in which for each one fL

reduction in MCV, the risk of GI cancer increased by 3%. Confirming MCV as a new predictor for GI cancer risk was consistent with other studies such as Capurso study (Capurso et al. 2004). Checking the model accuracy and performance, by validating it internally (objective 2), confirmed that age, sex, Hb, and MCV were all independent predictors of the risk of GI cancer.

Results showed also that FIT positivity is significantly associated with GI malignancy in IDA (objective 3). The finding that FIT has some predictive value for GI cancer in the IDA population is consistent with previous studies (Nakama et al. 2000; Cilona et al. 2011).

Limitations of this study include the small numbers of FIT participants. This resulted in a very low sensitivity of FIT for GI cancer with very wide confidence interval. Furthermore, because both transferrin saturation and serum ferritin values were only available in about one-third of the study population, the analysis of iron deficiency as an independent predictor of GI cancer was suboptimal in this study. Finally, using secondary data for the analysis meant that there was no opportunity to examine the predictive value of other important variables such as BMI, weight, etc.

Chapter 5. Broad external validation of a multivariable risk prediction model for GI malignancy in iron deficiency anaemia

This study (Almilaji et al. 2021b) externally validated the IDIOM model and stratified IDA patients into 5 risk groups based on the PPV, and NPV values. Externally validating the risk prediction model was a necessary step for this model to be used in practice.

The evaluation of the IDIOM model was carried out in this chapter by using not just one but two independent datasets from Oxford (n=1117) and Sheffield (n=474) (objective 4).

The findings of the external validation showed an acceptable discrimination performance of the model (about 70% in each dataset), no tendency for the IDIOM model to under or over-estimate of the cancer risk in both datasets, and a net benefit above the “investigate all” and “investigate no-one” strategies. The external validation analysis showed also that the IDIOM model can be applied with confidence to different populations in the UK.

The IDA patients were stratified into 5 GI risk groups (objective 5). Using a risk cut-off of 1.25% to categorise patients into the ultra-low risk group showed that none of the IDA patients within this group proved to have GI cancer on investigation in any dataset (Dorset, Oxford, and Sheffield).

As a retrospective analysis, limitations include the inability to control the size of the study external validation datasets, especially in the case of Sheffield dataset in which the number of events was small (36 events only). This has resulted in a restricted evaluation per centre.

Chapter 6. Modelling the episodes of care for IDA patients in a secondary-care centre using continuous-time multistate Markov chain

This study (Almilaji et al. 2022a) proposed a baseline model to represent episodes of care for IDA patients at a secondary care centre using continuous-time multi-state Markov chain (objective 6).

Preliminary results from this study showed that being diagnosed with positive GI cancer is always preceded by the prediction of being considered “very high risk” by the IDIOM model at the earlier visit.

Thus, the study raises the awareness of the importance of completing the GI investigations especially in IDA patients who are at high risk of GI cancer and

fit enough to do the investigations. Also, the proposed methods in this study could help policy makers in the future to establish what is the maximum delay time a confirmed IDA patient should be allowed to stay in before investigated, and what are the measures that could be put in place to reduce this time. However, due to the small sample size, preliminary results from this study, still need further validation.

Chapter 7. The development of a web-based application to predict the risk of gastrointestinal cancer in iron deficiency anaemia; the IDIOM App

This study (Almilaji et al. 2022b) documented in full detail the development and certifying process of the standalone software medical device app -the *IDIOM App*- in an academic setting by a PhD student.

Using R language, the *IDIOM App* Version 1.0, was successfully developed, registered with the MHRA and lawfully placed on the market with expected service life up to June 2023 under the present certification (objective 7).

There are different examples of published research that relates to developing apps in medical and healthcare setting (Andersson et al. 2020; Hsiao et al. 2006; Saho et al. 2021). However, this study may be proved to be of high value for other researchers who want to conduct a similar development project in the future as the resources for developing *medical device* apps are currently very limited.

A limitation of this chapter relates to the fact that the process of developing and certifying the *IDIOM App* was carried out according to the MDD regulations which were subsequently replaced by the more-strict MDR and more recently by the UKCA regulations. Consequently, some steps/details

that were described in this study may be different and not relevant to the new regulations.

Chapter 8. Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia

This study (Almilaji et al. 2021a) confirmed previous observations that IDA is strongly associated with right-sided CRC.

However, in contrast to other studies that suggest IDA is a marker of poor prognosis in CRC (Alexiusdottir et al. 2012; Tokunaga et al. 2019; Stapley et al 2006; Wilson et at. 2017), this study showed that diagnosis through an IDA pathway has the potential to downstage the disease (objective 8). Causes for these apparent dissimilarities might be because right-sided CRC is associated with diagnosis at a more advanced stage, and thus poorer prognosis (Mik et al. 2017; Yahagi et al. 2015; Petrelli et al. 2017; Snaebjomsson et al 2010; Hasen and Jess 2012; Nawa et al. 2008; Stebbing and Nath 1995; Goodman and Irvin 1993; Teng et al. 2014).

The study also examined the association between prior blood count test event and CRC stage (objective 9). Results showed that at any age, the risk of detecting CRC at later stage was higher in those without a previous blood count check. This observation may be because individuals who avoid medical care are inherently less likely to have a blood test, and thus less likely to present early with their undiagnosed CRC.

Limitations of this study include the uncertain applicability of a single centre experience in Dorset to other populations, and being a retrospective analysis, the inability to control the size of the study subgroups or to incorporate other variables such as grade, and MCV.

Chapter 9. Colorectal cancer and the blood loss paradox

This study (Almilaji et al. 2021c) further supported the recommendation of a blood count test in the at-high risk population from the former chapter.

The study compared the clinical characteristics of patients with CRC diagnosed via BCSP and IDA pathways and examined whether these pathways indeed identify different CRC sub-populations (objective 10).

The results from this chapter showed significant differences in the proportion with right-sided cancer and anaemia for patients with CRC diagnosed via FOB and IDA pathways. The implication is that these pathways identify distinct CRC sub-populations.

10.3 Significance of the study to the clinical practice

10.3.1 Introducing a unique threshold to refer IDA patients for suspected GI cancer pathway

In NICE guidelines (2015), PPV is used to determine the risk threshold, whereby if the risk of symptoms being caused by any type of cancer is above this threshold, then investigation or referral is warranted. Considering the financial implications, the guideline development group agreed to use a 3-5% PPV threshold value(s) to underpin the recommendations for suspected cancer pathway referrals and urgent direct access investigations, such as endoscopy. In the same recommendations, only smoking and age were found to be significantly predictors of cancer and hence should be included where relevant (NICE 2015).

The study showed that beside age; sex, Hb, and MCV are also significant predictors of GI cancer in IDA patients and hence all should be used when stratifying IDA patients in GI cancer risk groups.

Using all these significant predictors, the stratification of the IDA patients into GI risk groups was carried out in this study, based on the PPV quartiles, in which the lowest quarter is divided into two categories based on a risk threshold at which the NPV remains 100%.

The study showed that the lowest PPV value for the risk of GI cancer was about 8% in IDA patients, and which is well above the NICE PPV values of 3-5%. Thus, for suspected GI cancer pathway referrals in IDA population, the current NICE PPV values of 3-5% could lead to over-investigation of a large population with poor yield.

10.3.2 Stratifying IDA patients in GI cancer groups to prioritise investigations

Instead of classifying all patients with suspected cancer into two categories only (refer: above the 5% PPV value, or not to refer: below 5% PPV value), this study provides an opportunity to prioritise investigations especially when resources are limited by classifying IDA patients into five GI cancer risk groups. Therefore, patients who are classified in the highest risk group could be investigated as a priority, while patients who are classified in the second highest risk group could be investigated next, and so on.

In practice, stratifying IDA patients into high/very high-risk groups of GI cancer might lead to speeding the GI investigation and thus reducing its potential prognostic implications. Stratifying IDA patients into ultra-low risk group of GI cancer and by using also other safety net measures such as FIT, might lead to saving unnecessary referrals and investigations cost and helping patients to avoid invasive procedures.

10.3.3 Proposing blood count checks for at-risk population as a screening complementary element to the current Bowel Cancer Screening Programme in the UK

Colorectal cancer is one of the most common cancers and is currently one of the major causes of cancer-related mortality worldwide. This high mortality of CRC has made early diagnosis of CRC an important global issue (Han et al. 2019). Diagnosing cancer at early stages, often allows for more effective treatment options, and longer survival periods. To detect CRC as early as possible, many countries have population-based CRC screening programmes (Carroll et al. 2014). However, current CRC screening programmes have lower uptake than other cancer screening programmes, and lower sensitivity.

The uptake of CRC screening programme remains lower than that for other screening programmes (Koo et al. 2017). Only 50% to 58% of people in the UK who are invited for bowel cancer screening are screened adequately within 6 months of invitation (PHE 2015). In the routine breast screening programme, the level of the uptake never fell lower than 70% - the acceptable level of uptake- over the past 10 years period (National Health Service (NHS) Digital 2019). 71.4% is the uptake level in the cervical screening programme in 2017-2018 (National Health Service (NHS) Digital 2018).

Acceptability of the screening program may affect its uptake (Koo et al. 2017). For example, though colonoscopy is generally safe, it is still an invasive procedure with a 0.2% rate of severe complications (Nelson et al. 2002). With a 10-fold higher complications risk than any other commonly used cancer screening test (Ransohoff et al. 2009), it obviously will not be accepted by a substantial proportion of people.

Another factor affecting this low uptake may be how the bowel cancer screening programme is perceived by the screened population. Higher perceived “ease of completion” is found to be a significant predictor to complete the FIT test (Chambers et al. 2016a). On the contrary, many patients

were put off by their perception that completing the FOB test is a disgusting thing to do (Chambers et al. 2016b).

The low sensitivity of the CRC screening programme may be attributed to the inherited low sensitivity in the faecal tests, and the inefficiency of detecting right-side CRCs.

If the cancer was not bleeding when the screening faecal test was taken, there is a chance that a cancer might be missed (Public Health England 2019). At cut-off of 10 μg/g, FIT sensitivity estimated to be 92.1% (95% CI: 86.9% - 95.3%) (D'Souza et al. 2019). Although FIT sensitivity is better than FOB test, there are still 8% of people with CRC will be missed in this test. Moreover, FIT sensitivity is still lower than that of the other national screening programmes. The sensitivity of digital screening mammography in breast is 97% (Zeeshan et al. 2018). In a cervical screening programme, sensitivity of detecting CIN3 cervical cancer precursors in a colposcopy reached 97.2% (Wentzensen et al. 2012) and p16/Ki-67 dual-stained cytology showed a sensitivity of 100% (Uijterwaal et al. 2014).

Furthermore, in the UK, 74% of the screen-detected CRCs are found in the left side (Braun et al. 2016). And only 26% of these screen-detected CRCs are found in the right (Braun et al. 2016). This may be due to different reasons including often lower quality of cleansing of the right colon (Brenner et al. 2010). Also, because of their flat morphology, right side CRCs are much more difficult to be picked up than left side CRCs (Heresbach et al. 2008). Unfortunately, as a result, right-sided CRCs are commonly detected in more advanced stages than left side CRCs (Baran et al. 2018).

This study has shown clearly that CRC diagnosis through the identification of IDA pathway has the potential to downstage CRC- especially the right sided cases-, and at any age, the probability that CRC is detected at an early-stage is higher in those with a previous blood count.

Accordingly, the current approach to the diagnosis of CRC could be improved by complementing the existing BCSP by new blood count checks for IDA. These blood count checks have an advantage over the current CRC screening modalities of stool testing or sigmoidoscopy in being more acceptable to many people, have low cost, and more sensitive in detecting the right-sided CRCs. Indeed, all the WHO principles of early disease detection screening test (Wilson et al. 1968) are met in terms of the condition sought should be an important health problem, being acceptable test, easy to implement, lower cost, and effective.

Recommending the introduction of blood count checks for the at-risk population is in line with other studies that have tried to prioritise of colonoscopy/investigations in anaemia by identifying individuals at increased risk for CRC by analysing blood counts. Examples of these studies are Thompson et al. 2017; Kinar et al. 2016.

10.3.4 Introducing an automated tool to inform investigations decisions in a secondary care setting

To date there is no consensus on the risk threshold warranting investigation for GI cancer in IDA patients. Health professionals are relying on the discussion with their patients, and on their experience when estimating the GI cancer risk.

The IDIOM App is of potential use because it provides an immediate non-subjective reliable estimate for the GI risk in IDA, stratifies patients in meaningful risk groups, and inform the decisions of healthcare professionals and patients with regards to the future management of IDA patients, and appropriateness and timing of the investigations.

The *IDIOM App* represent the first ever software medical device that is developed, certified, and used in secondary care in the UK as clinical decision

support tool to estimate the risk of GI cancer in IDA. It has proven to be popular as the number of its users has been increasing steadily since it became put in service on 1st Dec 2020. The total up-to-date number of the app users is 762, and they come mainly from the UK (78%), then Spain (4%), India (2%), and from 45 countries around the world.

10.4 Academic and societal impacts

This PhD project involved the analysis of the *largest* dataset used ever to examine and internally validate the association between IDA and GI cancer. And represents the *first study* to externally validate a multivariable predictive model for GI risk in IDA patients by using more than one external dataset; stratify IDA patients into five risk groups based on the PPV, and NPV values; demonstrate and discuss the appropriate methods to model the IDA patients' episodes of care at a secondary-care centre; develop, certify, and document in full details the development of standalone software medical device app in an academic setting; and investigate the association between the event of having prior blood count check and the CRC stage in IDA patients.

The academic impact of the research includes the dissemination of research findings through various presentations at different international / national conferences and specialty academic meetings. These conferences and meetings were valuable opportunities to publish different outputs of this PhD project; and enabled the interaction and learning from experts in the domain, and the development of potential academic collaborations with researchers with similar research plans.

The academic impact of the research includes also the publications of original research articles with the PhD student as first author in peer-reviewed scientific journals from highly respected academic publishing groups such as the BMJ, Nature, etc. These articles were cited by other researchers during

the PhD project. One notable academic impact is the endorsement of the IDIOM App by the British Society of Gastroenterology guidelines for the management of iron deficiency anaemia in adults (Snook et al. 2021). This endorsement represents the pinnacle of this PhD academic impact. Because this clinical practice guidelines represents the highest level of evidence-based research ever in the field, that prepared by a team of experts to assist practitioners in making patient decisions and offer an evaluation of the quality of the relevant scientific literature.

The societal impact of the PhD research involves the usage of the IDIOM App exceptionally during the Covid-19 pandemic, as mentioned before in the introduction, when investigational resources were limited to assist with the triage of subjects with IDA in the Gastroenterology department at Poole hospital. The trial proved extremely successful, and therefore continues to be routinely used to assist with counselling of subjects with IDA since the restrictions have eased. The robustness of the app's model, the ease of accessing the app from different devices at any time, and the speed of displaying and explaining the GI cancer risk in a lay language made the app a very useful, popular, and reliable tool that fits well with the health professionals knowledge and experience when informing their discussions with regards to the future management of IDA patients.

10.5 Conclusion and future work

In conclusion, this study has confirmed and extended previous observations, showing that the simple and objective criteria of age, sex, Hb and MCV are strong and independent predictors of the risk of underlying GI cancer in subjects with IDA using the largest sample size used yet to study the association between GI cancer and IDA. The IDIOM model was validated internally and externally and showed a good promise of generalisability and transportability; and stratified IDA patients into different risk groups that could lead to fast track the investigation or save cost and help patients avoid

invasive procedures. The study highlights the importance of completing GI investigations and proposes a strong case for formally recommending the introduction of blood count checks of the at-risk population as a screening test complementary to the current BCSP. An automated decision-support tool, the *IDIOM App*, is developed in which anonymised, individual, patient data is entered, and GI cancer risk is calculated based on the IDIOM model and displayed immediately.

Subject to being successful in securing further funding, future plans that relate to different components of this PhD study include certifying the new version of the App pursuant to the new UKCA and MDR regulations after externally validating the IDIOM model by using new clinical datasets for patients from outside the UK and preferably from populations that do not share the same western diet such as India, China, etc. Attaining all the ethical approvals for adding a functionality to collect real-time clinical data from the IDIOM app is planned. Finally, validating the app in primary care setting as a decision-support tool to refer patients to secondary care similar to the usage of that information technology for identification of suspected colorectal cancer in primary care in the CREDIBLE study (Kidney et al. 2015).

New interesting research questions related to this PhD project, can be answered in the future by applying different statistical and machine learning methods. These questions relate to examining whether other factors such as ethnicity and BMI can be strong independent predictors of the risk of GI cancer in IDA. Estimating the appropriate age range and test frequency for the proposed IDA blood screening test. Assessing the clinical impact of the proposed IDA blood screening test on the available endoscopy services. Examining the causation between IDA and GI cancer and addressing the issue of whether IDA is a trigger of the GI cancer initiation or an outcome (symptom/sign) that is resulted from the cancer development. Finally, comparing the IDIOM model performance which is built using logistic regression to other artificial intelligence methods such as random forest and support vector machine.

References

Alexiusdottir, K. K., Möller, P. H., Snaebjornsson, P., Jonasson, L., Olafsdottir, E. J., Björnsson, E. S., Tryggvadottir, L., & Jonasson, J. G. 2012. Association of Symptoms of Colon Cancer Patients with Tumor Location and TNM Tumor Stage. Scandinavian Journal of Gastroenterology, 47(7), 795–801. doi:10.3109/00365521.2012.672589.

Altman, D. G., & Royston, P. 2006. The cost of dichotomising continuous variables. BMJ (Clinical research ed.), 332(7549), 1080. https://doi.org/10.1136/bmj.332.7549.1080.

Almilaji, O., Smith, C., Surgenor, S., Clegg, A., Williams, E., Thomas, P., & Snook, J. 2020. Refinement and Validation of the IDIOM Score for Predicting the Risk of Gastrointestinal Cancer in Iron Deficiency Anaemia. BMJ Open Gastroenterology.7: E000403. doi:10.1136/bmjgast-2020-000403.

Almilaji, O., Parry, S., Docherty, S., & Snook, J. 2021. Evidence for Improved Prognosis of Colorectal Cancer Diagnosed Following the Detection of Iron Deficiency Anaemia. Scientific Reports 11. 13055. doi:10.1038/S41598-021-92623-Z.

Almilaji, O., Webb, G., Maynard, A., Chapman, T. P., Shine, B., Ellis, A. J., Hebden, J., Docherty, S., Williams, E. J., & Snook, J. 2021. Broad external validation of a multivariable risk prediction model for gastrointestinal malignancy in iron deficiency anaemia. Diagnostic and prognostic research, 5(1), 23. https://doi.org/10.1186/s41512-021-00112-8.

Almilaji, O., Parry, S., Docherty, S., & Snook, J. 2021 Colorectal Cancer and the Blood Loss Paradox. Frontline Gastroenterology. Published Online First: 21 October 2021. doi: 10.1136/flgastro-2021-101959.

Almilaji, O. 2022. Modelling the Episodes of Care for IDA Patients in A Secondary-Care Centre Using Continuous-Time Multistate Markov Chain. Saudi J Gastroenterol. 28:115-21. DOI: 10.4103/sjg.sjg_387_21.

Almilaji, O., Engen, V., Snook, J., & Docherty, S. 2022. The development of a web-based application to predict the risk of gastrointestinal cancer in iron deficiency anaemia- the IDIOM app. Digital 2022, 2(1), 104-119; https://doi.org/10.3390/digital2010007.

American Joint Committee for Cancer Staging (AJCC) and End Results Reporting. 1977. Manual for Staging of Cancer (1st ed). Chicago, IL.

American Cancer Society. 2020. Treatment of Colon cancer. Available from: https://www.cancer.org/cancer/colon-rectal-cancer/treating/by-stage-colon.html [Accessed 8 Nov 2021].

Andersson, H. B., BØrresen, T., Prasolova-FØrland, E., McCallum, S., & Estrada, J. G. 2020. Developing An AR Application For Neurosurgical Training: Lessons Learned For Medical Specialist Education. IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pp. 407-412, doi: 10.1109/VRW50115.2020.00087.

Ayling, R. M., Lewis, S. J., & Cotter, F. (2019). Potential roles of artificial intelligence learning and faecal immunochemical testing for prioritisation of colonoscopy in anaemia. British journal of haematology, 185(2), 311–316. https://doi.org/10.1111/bjh.15776.

Baran, B., Mert Ozupek, N., Yerli Tetik, N., Acar, E., Bekcioglu, O., & Baskin, Y. 2018. Difference Between Left-Sided and Right-Sided Colorectal Cancer: A Focused Review of Literature. Gastroenterology Research, 11(4), 264–273. doi:10.14740/gr1062w.

Baxter, N. N., Goldwasser, M. A., Paszat, L. F., Saskin, R., Urbach, D. R., & Rabeneck, L. (2009). Association of colonoscopy and death from colorectal cancer. Annals of internal medicine, 150(1), 1–8. https://doi.org/10.7326/0003-4819-150-1-200901060-00306.

Beg, S., Ragunath, K., Wyman, A., Banks, M., Trudgill, N., Pritchard, D. M., Riley, S., Anderson, J., Griffiths, H., Bhandari, P., Kaye, P., & Veitch, A. (2017). Quality standards in upper gastrointestinal endoscopy: a position statement of the British Society of Gastroenterology (BSG) and Association of Upper Gastrointestinal Surgeons of Great Britain and Ireland (AUGIS). Gut, 66(11), 1886–1899. https://doi.org/10.1136/gutjnl-2017-314109.

Braun, M., Hill, J., Kuryba, A., Roe, A., Vallance, A., vanderMeulen, J., & Walker, K. 2016. National Bowel Cancer Audit (2ed). Available from https://www.hqip.org.uk/wp-content/uploads/2018/02/national-bowel-cancer-audit-annual-report-2016.pdf [Accessed 4 Oct 2021].

Brenner, H., Hoffmeister, M., Arndt, V., Stegmaier, C., Altenhofen, L., & Haug, U. 2010. Protection from Right-and Left-Sided Colorectal Neoplasms After Colonoscopy: Population-Based Study. Journal of the National Cancer Institute, 102(2), 89-95. doi:10.1093/Jnci/djp436.

Bufill, J. A. 1990. Colorectal Cancer: Evidence for Distinct Genetic Categories Based on Proximal or Distal Tumor Location. Annals of Internal Medicine, 113(10), 779–788. doi:10.7326/0003-4819-113-10-779.

Busti, F., Marchi, G., Ugolini, S., Castagna, A., & Girelli, D. (2018). Anemia and Iron Deficiency in Cancer Patients: Role of Iron Replacement Therapy. *Pharmaceuticals (Basel, Switzerland)*, *11*(4), 94. https://doi.org/10.3390/ph11040094.

Camaschella, C. 2017. New Insights into Iron Deficiency and Iron Deficiency Anemia. Blood Reviews, 31(4), 225–233. doi:10.1016/j.blre.2017.02.004.

Capurso, G., Baccini, F., & Osborn, J. 2004. Can patient characteristics predict the outcome of endoscopic evaluation of iron deficiency anemia: a multiple logistic regression analysis. Gastrointestinal Endoscopy. 59:766-771.

Carroll MR, Seaman HE, Halloran SP. Tests and investigations for colorectal cancer screening. Clin Biochem. 2014;47:921–39.

Chambers, J. A, O'Carroll, R. E., & Brownlee, L. 2016. Colorectal Cancer Screening and Perceived Disgust: The Importance of the “Ick” Factor in Faecal Occult Blood Test Uptake. Colorec Cancer. 2: 1. doi:10.21767/2471-9943.100013.

Chambers, J. A., Callander, A. S., Grangeret, R., & O'Carroll, R. E. 2016. Attitudes towards the Faecal Occult Blood Test (FOBT) Versus the Faecal Immunochemical Test (FIT) for Colorectal Cancer Screening: Perceived Ease of Completion and Disgust. BMC Cancer, 16, 96. doi:10.1186/S12885-016-2133-4.

Chowdhury, A., Longcroft-Wheaton, G., Davis, A., Massey, D., & Goggin, P. 2014. Role of Faecal Occult Bloods in the Diagnosis of Iron Deficiency Anaemia. Frontline Gastroenterology, 5(4), 231–236. doi:10.1136/flgastro-2013-100425.

Cilona, A., Zullo, A., Hassan, C., Ridola, L., & Annese, M. (2011). Is faecal-immunochemical test useful in patients with iron deficiency anaemia and without overt bleeding? Digestive and liver disease: official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver, 43(12), 1022–1024. https://doi.org/10.1016/j.dld.2011.08.002.

Conrad, M. E., Weintraub, L. R., Sears, D. A., & Crosby, W. H. 1966. Absorption of Hemoglobin Iron. The American Journal of Physiology, 211(5), 1123–1130. doi:10.1152/ajplegacy.1966.211.5.1123.

D'Souza, N., Brzezicki, A., & Abulafi, M. 2019. Faecal Immunochemical Testing in General Practice. The British Journal of General Practice: The Journal of the Royal College of General Practitioners, 69(679), 60–61. doi:10.3399/bjgp19x700853.

De Franceschi, L., Iolascon, A., Taher, A., & Cappellini, M. D. 2017. Clinical Management of Iron Deficiency Anemia in Adults: Systemic Review on Advances in Diagnosis and Treatment. European Journal of Internal Medicine, 42, 16–23. doi:10.1016/j.ejim.2017.04.018.

Disease and Injury Incidence and Prevalence Collaborators. 2016. Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 310 Diseases and Injuries, 1990-2015: A Systematic Analysis for the Global Burden of Disease Study 2015. Lancet. Oct 8;388 10053: 1545-1602.

Dukes, C. E. 1932. The Classification of Cancer of the Rectum. Journal of Pathological Bacteriology 35(3): 323-332.

Edna, T. H., Karlsen, V., Jullumstrø, E., & Lydersen, S. 2012. Prevalence of Anaemia at Diagnosis of Colorectal Cancer: Assessment of Associated Risk Factors. Hepato-Gastroenterology, 59(115), 713–716. doi:10.5754/hge11479.

Gervaz, P., Bucher, P., & Morel, P. 2004. Two Colons-Two Cancers: Paradigm Shift and Clinical Implications. Journal of Surgical Oncology, 88(4), 261–266. doi:10.1002/jso.20156.

Goddard, A. F., James, M. W., Mcintyre, A. S., Scott, B. B., & British Society of Gastroenterology 2011. Guidelines for the Management of Iron Deficiency Anaemia. Gut, 60(10), 1309–1316. doi:10.1136/gut.2010.228874.

Goodman, D., & Irvin, T. T. 1993. Delay in the Diagnosis and Prognosis of Carcinoma of the Right Colon. The British Journal of Surgery, 80(10), 1327–1329. doi:10.1002/bjs.1800801037.

Gualco, G., Reissenweber, N., Cliché, I., & Bacchi, C. E. (2006). Flat elevated lesions of the colon and rectum: a spectrum of neoplastic and nonneoplastic entities. *Annals of diagnostic pathology*, *10*(6), 333–338. https://doi.org/10.1016/j.anndiagpath.2006.03.003.

Han, Y.D., Oh, T.J., Chung, TH. 2019. Early detection of colorectal cancer based on presence of methylated syndecan-2 (SDC2) in stool DNA. Clin Epigenet 11, 51. https://doi.org/10.1186/s13148-019-0642-0.

Hansen, I., Jess, P. 2012. Possible Better Long-Term Survival in Left Versus Right-Sided Colon Cancer – A Systematic Review. Danish Medical Journal. 59: A4444.

Hawkes, N. 2019. Cancer survival data emphasise importance of early diagnosis. BMJ 364 :l408 doi:10.1136/bmj.l408.

Heresbach, D., Barrioz, T., Lapalus, M. G., Coumaros, D., Bauret, P., Potier, P., Sautereau, D., Boustière, C., Grimaud, J. C., Barthélémy, C., Sée, J., Serraj, I., D'Halluin, P. N., Branger, B., & Ponchon, T. 2008. Miss Rate for Colorectal Neoplastic Polyps: A Prospective Multicenter Study of Back-to-Back Video Colonoscopies. Endoscopy, 40(4), 284–290. doi:10.1055/S-2007-995618.

Hiom, S. C. 2015. Diagnosing Cancer Earlier: Reviewing the Evidence for Improving Cancer Survival. British Journal of Cancer, 112 Suppl 1(Suppl 1), S1–S5. doi:10.1038/bjc.2015.23.

Ho, C. H., Chau, W. K., Hsu, H. C. 2005. Predictive risk factors and prevalence of malignancy in patients with iron deficiency anemia in Taiwan. American Journal of Hematology. 78:108-112.

Hsiao, C. H., Hsu, T. C., Chang, J. N., Yang, S. J., Young, S. T., & Chu, W. C. 2006. Developing a medical image content repository for e-learning. Journal of digital imaging, 19(3), 207–215. https://doi.org/10.1007/s10278-006-0588-6.

Hung, N., Shen, C. C., Hu, Y. W., Hu, L. Y., Yeh, C. M., Teng, C. J., Kuan, A. S., Chen, S. C., Chen, T. J., & Liu, C. J. 2015. Risk of Cancer in Patients with Iron Deficiency Anemia: A Nationwide Population-Based Study. Plos One, 10(3), E0119647. doi:10.1371/journal.pone.0119647.

Hunt, G. C., & Faigel, D. O. (2002). Endoscopic evaluation of patients with partial gastrectomy and iron deficiency. Digestive diseases and sciences, 47(3), 641–644. https://doi.org/10.1023/a:1017988607854.

Iacopetta, B. 2002. Are there Two Sides to Colorectal Cancer? International Journal of Cancer, 101(5), 403–408. doi:10.1002/ljc.10635.

Imperial, R., Ahmed, Z., Toor, O. M., Erdoğan, C., Khaliq, A., Case, P., Case, J., Kennedy, K., Cummings, L. S., Melton, N., Raza, S., Diri, B., Mohammad, R., El-Rayes, B., Pluard, T., Hussain, A., Subramanian, J., & Masood, A. 2018. Comparative Proteogenomic Analysis of Right-Sided Colon Cancer, Left-Sided Colon Cancer and Rectal Cancer Reveals Distinct Mutational Profiles. Molecular Cancer, 17(1), 177. doi:10.1186/S12943-018-0923-9.

Ioannou, G. N., Rockey, D. C., Bryson, C. L., & Weiss, N. S. (2002). Iron deficiency and gastrointestinal malignancy: a population-based cohort study. The American journal of medicine, 113(4), 276–280. https://doi.org/10.1016/s0002-9343(02)01214-7.

James, M. W., Chen, C. M., Goddard, W. P. 2005. Risk factors for gastrointestinal malignancy in patients with iron-deficiency anaemia. European Journal of Gastroenterology & Hepatology. 17:1197-1203.

Johnson-Wimbley, T. D., & Graham, D. Y. 2011. Diagnosis and Management of Iron Deficiency Anemia in the 21st Century. Therapeutic Advances in Gastroenterology, 4(3), 177–184. doi:10.1177/1756283X11398736.

Johnstone, C., & Rich, S. E. (2018). Bleeding in cancer patients and its treatment: a review. Annals of palliative medicine, 7(2), 265–273. https://doi.org/10.21037/apm.2017.11.01.

Kidney, E., Berkman, L., Macherianakis, A., Morton, D., Dowswell, G., Hamilton, W., Ryan, R., Awbery, H., Greenfield, S., & Marshall, T. (2015). Preliminary results of a feasibility study of the use of information technology for identification of suspected colorectal cancer in primary care: the CREDIBLE study. British journal of cancer, 112 Suppl 1(Suppl 1), S70–S76. https://doi.org/10.1038/bjc.2015.45.

Kim, N. H., Lee, M. Y., Park, J. H., Park, D. I., Sohn, C. I., Choi, K., & Jung, Y. S. (2017). A Combination of Fecal Immunochemical Test Results and Iron Deficiency Anemia for Detection of Advanced Colorectal Neoplasia in Asymptomatic Men. Yonsei medical journal, 58(5), 910–917. https://doi.org/10.3349.

Kinar, Y., Kalkstein, N., Akiva, P., Levin, B., Half, E. E., Goldshtein, I., Chodick, G., & Shalev, V. (2016). Development and validation of a predictive model for detection of colorectal cancer in primary care by analysis of complete blood counts: a binational retrospective study. Journal of the American Medical Informatics Association: JAMIA, 23(5), 879–890. https://doi.org/10.1093/jamia/ocv195.

Koo, S., Neilson, L. J., Von Wagner, C., & Rees, C. J. 2017. The NHS Bowel Cancer Screening Program: Current Perspectives on Strategies for Improvement. Risk Management and Healthcare Policy, 10, 177–187. doi:10.2147/RMHP.S109116.

Lee, M. W., Pourmorady, J. S., & Laine, L. 2020. Use of fecal occult blood testing as a diagnostic tool for clinical indications: a systematic review and meta-analysis. Am J Gastroenterol. 115:662-670. doi:10.14309/ajg.0000000000000495.

Logan, R. F., Patnick, J., Nickerson, C., Coleman, L., Rutter, M. D., Von Wagner, C., & English Bowel Cancer Screening Evaluation Committee. 2012. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England After the First 1 Million Tests. Gut, 61(10), 1439–1446. doi:10.1136/gutjnl-2011-300843.

Lopez, A., Cacoub, P., Macdougall, I. C., & Peyrin-Biroulet, L. 2016. Iron Deficiency Anaemia. Lancet (London, England), 387(10021), 907–916. doi:10.1016/s0140-6736(15)60865-0.

Ludwig, H., Van Belle, S., Barrett-Lee, P., Birgegård, G., Bokemeyer, C., Gascón, P., Kosmidis, P., Krzakowski, M., Nortier, J., Olmi, P., Schneider, M., & Schrijvers, D. 2004. The European Cancer Anaemia Survey (ECAS): A Large, Multinational, Prospective Survey Defining the Prevalence, Incidence, and Treatment of Anaemia in Cancer Patients. European Journal of Cancer (Oxford, England: 1990), 40(15), 2293–2306. doi:10.1016/J.Ejca.2004.06.019.

Lynch, J. P., & Rustgi, A. K. 2012. Chapter 80 - Mechanisms of GI Malignancies, Editor S: Leonard R. Johnson, Fayez K. Ghishan, Jonathan D. Kaunitz, Juanita L. Merchant, Hamid M. Said, Jackie D. Wood, Physiology of the Gastrointestinal Tract (5th ed). Academic Press. 2129-2155. ISBN 9780123820266.

Mankodi, S., Hayee, B. H., O'Donohue, J., & Reffitt, D. (2010). Anaemia investigation in practice: inappropriate, cost inefficient with a risk of missing gastrointestinal cancer. Can we improve? Clinical medicine (London, England), 10(2), 115–118. https://doi.org/10.7861/clinmedicine.10-2-115.

Mik, M., Berut, M., Dziki, L., Trzcinski, R., & Dziki, A. 2017. Right- and Left-Sided Colon Cancer - Clinical and Pathological Differences of the Disease Entity in One Organ. Archives of Medical Science: AMS, 13(1), 157–162. doi:10.5114/Aoms.2016.58596.

Medical Device Directive (MDD) EU 93/42/EEC of 14 June 1993 concerning medical devices as amended by Directive 2007/47/EC.
Medical Device Regulation (MDR) EU 2017/745.

Nakama, H., Zhang, B., Kamijo, N., & Fattah, A. S. 2001. Frequency of Colorectal Cancer in Patients with Iron Deficiency Anaemia. Digestive and Liver Disease: official Journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver, 33(1), 88. doi:10.1016/S1590-8658(01)80145-6.

National Cancer Registration and Analysis Service (NCRAS). 2016. Routes to Diagnosis of Cancer by Stage. Available from http://www.ncin.org.uk/publications/routes_to_diagnosis [Accessed 4 Oct 2021].

National Institute for Health and Care Excellence (NICE). 2015. Suspected Cancer: Recognition and Referral. Last Updated: 29 January 2021. Available from: https://www.nice.org.uk/Guidance/Ng12 [Accessed 4 Oct 2021].

National Institute for Health and Care Excellence (NICE). 2021. Anaemia-Iron Deficiency. Available from: https://cks.nice.org.uk/topics/Anaemia-Iron-Deficiency [Accessed 4 Oct 2021].

Nawa, T., Kato, J., Kawamoto, H., et al. 2008. Differences Between Right and Left-Sided Colon Cancer in Patient Characteristics, Cancer Morphology and Histology. Journal of Gastroenterology and Hepatology. 23:418-23. https://doi.org/10.1111/j.1440-1746.2007.04923.x.

Nelson, D. B., Mcquaid, K. R., Bond, J. H., Lieberman, D. A., Weiss, D. G., & Johnston, T. K. 2002. Procedural Success and Complications of Large-Scale Screening Colonoscopy. Gastrointestinal Endoscopy, 55(3), 307–314. doi:10.1067/mge.2002.121883.

National Health Service (NHS) Digital. 2018. Cervical Screening Programme 2017-18. Available from https://digital.nhs.uk/data-and-information/publications/statistical/cervical-screening-annual/england---2017-18 [Accessed 4 Oct 2021].

National Health Service (NHS) Digital. 2019. Proportion of Women Taking Up Breast Screening Invitations Falls. Available from https://digital.nhs.uk/news/2019/proportion-of-women-taking-up-breast-screening-invitations-falls [Accessed 4 Oct 2021].

National Health Service (NHS). 2018. What Do Cancer Stages and Grades Mean? Available from https://www.nhs.uk/Common-Health-Questions/Operations-Tests-and Procedures/What-Do-Cancer-Stages-and-Grades-Mean/ [Accessed 4 Oct 2021].

National Health Service (NHS). 2019. Bowel Cancer. Available from https://www.nhs.uk/Conditions/Bowel-Cancer/Treatment/ [Accessed 4 Oct 2021].

National Health Service (NHS). 2019. Stomach Cancer. Available from https://www.nhs.uk/Conditions/Bowel-Cancer/Treatment/ [Accessed 4 Oct 2021].

Nitsche, U., Stogbauer, F., Spath, C., Haller, B., Wilhelm, D., Friess, H., Bader, F. G. Right sided colon cancer as a distinct histopathological subtype with reduced prognosis. Dig Surg. 2016;33(2):157–163. doi: 10.1159/000443644.

Niv, E., Elis, A., Zissin, R., Naftali, T., Novis, B., & Lishner, M. (2005). Iron deficiency anemia in patients without gastrointestinal symptoms--a prospective study. Family practice, 22(1), 58–61. https://doi.org/10.1093/fampra/cmh705.

Office for National Statistics (ONS). 2019. Cancer Survival by Stage at Diagnosis for England. Available from https://www.ons.gov.uk/Peoplepopulationandcommunity/Healthandsocialcare/Conditionsanddiseases/Bulletins/Cancersurvivalbystageatdiagnosisforenglandexperimentalstatistics [Accessed 4 Oct 2021].

Patel, S., Jose, A., & Mohiuddin, S. S. 2021. Physiology, Oxygen Transport and Carbon Dioxide Dissociation Curve. Statpearls Publishing, Treasure Island (FL).

Petrelli, F., Tomasello, G., Borgonovo, K., et al. 2017. Prognostic Survival Associated with Left-Sided Vs Right-Sided Colon Cancer: A Systematic Review and Meta-Analysis. JAMA Oncology. 3:211–19.

Polin, V., Coriat, R., Perkins, G., Dhooge, M., Abitbol, V., Leblanc, S., Prat, F., & Chaussade, S. 2013. Iron Deficiency: from Diagnosis to Treatment. Digestive and Liver Disease: official Journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver, 45(10), 803–809. doi:10.1016/j.dld.2013.02.019.

Public Health England (PHE). 2015. Bowel Screening, Data for FY 2014/15. Available from https://assets.publishing.service.gov.uk/government/uploads/system/upload

s/attachment_data/file/574716/Screening_in_England_2014_to_2015.pdf [Accessed 4 Oct 2021].

Public Health England (PHE). 2015. FOBt Screening, Data for FY 2014/15. Available fromhttps://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/574716/Screening_in_England_2014_to_2015.pdf [Accessed 4 Oct 2021].

Public Health England (PHE). 2017. The 2nd Atlas of Variation in NHS Diagnostic Services in England. Available from https://fingertips.phe.org.uk/Profile/Atlas-of-Variation. [Accessed 4 Oct 2021].
Public Health England (PHE). 2019. Bowel Cancer Screening Programme BCSP: Overview. Available from https://www.gov.uk/guidance/bowel-cancer-screening-programme-overview [Accessed 4 Oct 2021].

Raffin, S. B., Woo, C. H., Roost, K. T., Price, D. C., & Schmid, R. 1974. Intestinal Absorption of Hemoglobin Iron-Heme Cleavage by Mucosal Heme Oxygenase. The Journal of Clinical Investigation, 54(6), 1344–1352. doi:10.1172/JCI107881.

Rai, S., & Hemingway, D. (2005). Iron deficiency anaemia--useful diagnostic tool for right sided colon cancers? Colorectal disease: the official journal of the Association of Coloproctology of Great Britain and Ireland, 7(6), 588–590. https://doi.org/10.1111/j.1463-1318.2005.00880.x.

Ransohoff, D. F. 2009. How Much Does Colonoscopy Reduce Colon Cancer Mortality? Annals of Internal Medicine, 150(1), 50–52. doi:10.7326/0003-4819-150-1-200901060-00308.

Rao, J. N., & Wang, J. Y. 2010. Regulation of Gastrointestinal Mucosal Growth. San Rafael (CA): Morgan & Claypool Life Sciences.

Rockey, D. C., & Cello, J. P. (1993). Evaluation of the gastrointestinal tract in patients with iron-deficiency anemia. The New England journal of medicine, 329(23), 1691–1695. https://doi.org/10.1056/NEJM199312023292303.

Rubin, G., Walter, F., Emery, J., & De Wit, N. 2018. Reimagining the Diagnostic Pathway for Gastrointestinal Cancer. Nature Reviews. Gastroenterology & Hepatology, 15(3), 181–188. doi:10.1038/Nrgastro.2018.1.

Saho, N. J. G., & Ezin, E. C. 2021. Developing a Digital Healthcare Book based on the Blockchain Technology. International Symposium on Electrical, Electronics and Information Engineering (ISEEIE 2021). Association for Computing Machinery, New York, NY, USA, 617–622. DOI:https://doi.org/10.1145/3459104.3459203.

Schop, A., Stouten, K., Riedl, J., Van Houten, R., Van Rosmalen, J., Wolfhagen, F., Bindels, P., & Levin, M. D. 2019. Long-Term Outcomes in Patients Newly Diagnosed with Iron Deficiency Anaemia in General Practice: A Retrospective Cohort Study. BMJ Open, 9(11), E032930. doi:10.1136/bmjopen-2019-032930.

Selby, K., Levine, E. H., Doan, C., Gies, A., Brenner, H., Quesenberry, C., Lee, J. K., & Corley, D. A. (2019). Effect of Sex, Age, and Positivity Threshold on Fecal Immunochemical Test Accuracy: A Systematic Review and Meta-analysis. Gastroenterology, 157(6), 1494–1505. https://doi.org/10.1053/j.gastro.2019.08.023.

Shahriari-Ahmadi, A., Shalbaf, N., Masoodi, M., Shalbaf, M.,& Bozorgi, V. 2017. Endoscopic and Colonoscopic Findings in Patients with Iron Deficiency Anemia: The Risk of Cancer, Int J Cancer Manag. 2017; 10(10): E8222. doi:10.5812/Ijcm.8222.

Shaw, A. G., Simpson, J., Tierney, G., Goddard, A. F., Reynolds, J. R., & Lund, J. N. (2008). Referral of patients with iron deficiency anaemia under the lower gastrointestinal two-week wait rule. Colorectal disease: the official journal of the Association of Coloproctology of Great Britain and Ireland, 10(3), 294–297. https://doi.org/10.1111/j.1463-1318.2007.01364.x.

Silva, A. C., Sheppard, Z. A., Surgenor, S. L., Williams, E. J., Thomas, P. W., & Snook, J. A. 2014. Clinical Risk Factors for Underlying Gastrointestinal Malignancy in Iron Deficiency Anaemia: the IDIOM Study. Frontline Gastroenterology, 5(4), 237–242. doi:10.1136/flgastro-2013-100386.

Snaebjornsson, P., Jonasson, L., Jonsson, L., et al. 2010. Colon Cancer in Iceland: A Nationwide Comparative Study on Various Pathology Parameters with Respect to Right and Left Tumor Location and Patients Age. International Journal of Cancer. 127:2645-53.

Snook, J., Bhala, N., Beales, I., Cannings, D., Kightley, C., Logan, R. P., Pritchard, D. M., Sidhu, R., Surgenor, S., Thomas, W., Verma, A. M., & Goddard, A. F. 2021. British Society of Gastroenterology Guidelines for the Management of Iron Deficiency Anaemia in Adults. Gut, 70(11), 2030–2051. doi:10.1136/gutjnl-2021-325210.

Stapley, S., Sharp, D., & Hamilton, W. 2006. The Mortality of Colorectal Cancer in Relation to The Initial Symptom and To the Duration of Symptoms: A Cohort Study in Primary Care. British Journal of Cancer. 95:1321–25.

Stebbing, J. F., & Nash, A. G. 1995. Avoidable Delay in the Management of Carcinoma of the Right Colon. Annals of the Royal College of Surgeons of England, 77(1), 21–23.

Stone, H., Almilaji, O., John, C., Smith, C., Surgenor, S., Lachlan, A., Williams, E. J., & Snook, J. 2020. The Dedicated Iron Deficiency Anaemia

Clinic: A 15-Year Experiencefrontline Gastroenterology Published Online First: 10 September. doi:10.1136/flgastro-2020-101470.

Teng, C., L., J, Yu, J. T., Chen, Y. H. 2014. Early Colonoscopy Confers Survival Benefits on Colon Cancer Patients with Pre-Existing Iron Deficiency Anemia: A Nationwide Population-Based Study. PLoS ONE. 9(1): e86714. Available at: https://doi.org/10.1371/journal.pone.0086714.

Thompson, M. R., O'Leary, D. P., Flashman, K., Asiimwe, A., Ellis, B. G., & Senapati, A. (2017). Clinical assessment to determine the risk of bowel cancer using Symptoms, Age, Mass and Iron deficiency anaemia (SAMI). The British journal of surgery, 104(10), 1393–1404. https://doi.org/10.1002/bjs.10573.

Tian, C., & Hu. G. 2020. Chapter Twenty-Five - Chromatin Isolation by RNA. Purification Chirp and Its Applications, Editor S: Trygve tollefsbol, in Translational Epigenetics, Epigenetics Methods, Academic Press, 18: 507-521, doi:10.1016/b978-0-12-819414-0.00025-2.

Tokunaga, R., Nakagawa, S., Miyamoto, Y., Ohuchi, M., Izumi, D., Kosumi, K., Taki, K., Higashi, T., Miyata, T., Yoshida, N., & Baba, H. 2019. The Impact of Preoperative Anaemia and Anaemic Subtype on Patient Outcome in Colorectal Cancer. Colorectal Disease: the official Journal of the Association of Coloproctology of Great Britain and Ireland, 21(1), 100–109. doi:10.1111/codi.14425.

Treuting, P., Dintzis, S. M., & Montine K. S. New York, NY: Academic Press, 2017. ISBN-13: 978-0128029008.

Uijterwaal, M. H., Witte, B. I., Van Kemenade, F. J., Rijkaart, D., Ridder, R., Berkhof, J., Balfoort-Van Der Meij, G. A., Bleeker, M. C., Snijders, P. J., & Meijer, C. J. 2014. Triaging Borderline/Mild Dyskaryotic Pap Cytology with P16/Ki-67 Dual-Stained Cytology Testing: Cross-Sectional and Longitudinal

Outcome Study. British Journal of Cancer, 110(6), 1579–1586. doi:10.1038/bjc.2014.34.

Union for International Cancer Control (UICC). 2016. TNM Classification of Malignant Tumours (8th ed). Wiley-Blackwell. ISBN: 978-1-119-26357-9.

Wentzensen, N., Schwartz, L., Zuna, R. E., Smith, K., Mathews, C., Gold, M. A., Allen, R. A., Zhang, R., Dunn, S. T., Walker, J. L., & Schiffman, M. 2012. Performance of P16/Ki-67 Immunostaining to Detect Cervical Cancer Precursors in A Colposcopy Referral Population. Clinical Cancer Research: An official Journal of the American Association for Cancer Research, 18(15), 4154–4162. doi:10.1158/1078-0432.CCR-12-0270.

Wijayasekara, C. M., Surgenor,S. L., Hatcher, C., Clegg, A. J., Williams, E. J., & Snook, J.2016 Clinical Risk Factors for Underlying Gastrointestinal Malignancy in Iron Deficiency Anaemia–Prospective Validation of the IDIOM Score, Cogent Medicine, 3: 1, 1213216, doi:10.1080/2331205X.2016.1213216.

Wilson, M. J., van Haaren, M., Harlaar, J. J., 2017. Long-Term Prognostic Value of Preoperative Anemia In Patients with Colorectal Cancer: A Systematic Review and Meta-Analysis. Surgery Oncology. 26:96-104.

Wilson, J. M. G., Jungner, G. & World Health Organization. 1968. Principles and practice of screening for disease / J. M. G. Wilson, G. Jungner. World Health Organization. https://apps.who.int/iris/handle/10665/37650.

World Health Organization (WHO). 2011. Prevention of Iron Deficiency Anaemia in Adolescents. Geneva: United Nations University. Available from: https://apps.who.int/Iris/Bitstream/10665/205656/1/B4770.Pdf?Ua=1 [Accessed 4 Oct 2021].

World Health Organization (WHO). 2021. Cancer, Fact Sheet. Available from: https://www.who.int/News-Room/Fact-Sheets/Detail/Cancer [Accessed 4 Oct 2021].

Yahagi, M., Okabayashi, K., Hasegawa, H., Tsuruta, M., Kitagawa, Y. 2015. The Worse Prognosis of Right-Sided Compared with Left-Sided Colon Cancers: a Systematic Review and Meta-analysis. J Gastrointest Surg. 2016 Mar;20(3):648-55. doi: 10.1007/s11605-015-3026-6.

Zeeshan, M., Salam, B., Khalid, Q., Alam, S., & Sayani, R. 2018. Diagnostic Accuracy of Digital Mammography in the Detection of Breast Cancer. Cureus, 10(4), E2448. doi:10.7759/cureus.2448.

Zhu, A., Kaneshiro, M., & Kaunitz, J. D. 2010. Evaluation and Treatment of Iron Deficiency Anemia: A Gastroenterological Perspective. Digestive Diseases and Sciences, 55(3), 548–559. doi:10.1007/S10620-009-1108-6.

Appendices

Appendix I: Research outputs

Appendix Ia: Peer reviewed conference abstracts

Appendix Ia1: 10th Annual Postgraduate BU Conference. Mar 2018

Accepted abstract:

The Importance of Identifying Iron Deficiency Anaemia in the Early Detection of Colorectal Cancer

This is a 13-year retrospective analysis of data collected from 2295 patients who attended iron deficiency anaemia (IDA) outpatient clinic in Poole hospital. The datasets were anonymised and consist of five main groups, including demographic data, blood test results, investigations, outcome of the investigations, and cancer data. The aim of this study is to investigate the association between colorectal cancer (CRC) and IDA, and discover whether new risk factors can be identified and contributed to earlier diagnosis of CRC through the detection of IDA. R programming language will be used for the statistical analysis. Clustering methods will be performed to discover distinct patients' groups and patterns. And multinomial logistic regression will categorize patient profile (blood test results, age, and gender) as either low CRC risk, or high CRC risk. Decision trees and support vector machines [SVM], will be applied also and all will be compared to select the optimal method.

Presented poster:

The Importance of Identifying Iron Deficiency Anaemia in the Early Detection of Colorectal Cancer

Bournemouth University

O. Almilaji*, P. Thomas, J. Snook

*PhD Researcher, Bournemouth University, Faculty of Health and Social Sciences, Royal London House
E-mail:o.almilaji@bournemouth.ac.uk
Office: 313, Tel: 01202 966741

1. Background

Colorectal cancer (CRC) is the second most common cause of cancer deaths in the UK (1). The UK has its NHS bowel cancer screening programme that includes two elements: Faecal Occult Blood test (FOB) which checks for hidden blood in the stool, and **Bowel Scope** which examines the lower part of the bowel to detect CRCs and remove polyps that may advance into cancers.

A Bowel Scope Procedure to remove a polyp

FOB Test Kit

A potential currently neglected approach to the earlier diagnosis of CRC is through the detection of iron deficiency anaemia (IDA). It appears that IDA is a feature of a different subgroup of CRC, in particular most occur in the right colon (2). IDA can result from long-term blood loss from a gastrointestinal lesion (GI). In IDA, the red blood cells (RBC) are smaller (lower mean corpuscular volume (MCV)) and more pale (lower Haemoglobin (Hb) level). In severe IDA, Ferritin (i.e. total amount of iron stored in the body) and Transferrin saturation (i.e. how much serum iron is bound (T.satu)) are usually much lower than the normal levels.

Normal RBCs IDA' RBCs Bleeding CRC Tumour

Research Question: Could the usage of IDA as a complementary element to the current NHS bowel cancer screening programme enhance the earlier diagnosis of CRC?
Investigating the association between CRC and IDA, and identifying the subgroups of IDA patients who are at increased/lower risk of CRC might lead to speeding the investigation of CRC and thus reducing its potential prognostic implications, saving cost by avoiding over-investigation of a large population with poor yield, and helping patients avoid invasive procedures.

2. Aims

- Refine the current cut offs for CRC risk factors and assess the prevalence of IDA in patients with CRC.
- Assess the current frequency of blood count testing to detect IDA in at-risk age-groups.
- Assess whether cancers diagnosed through the detection of IDA are associated with an earlier tumour stage.
- Assess whether FOB testing can improve the clinical prediction of CRC in patients with IDA.

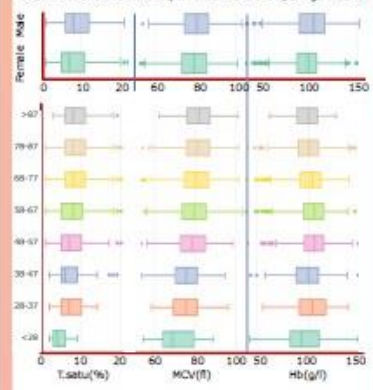
3. Data and Methods

Data: This is a 13-year (2004–2016) **retrospective cohort** analysis of data collected prospectively from **2295 patients** who attended the IDA outpatient clinic in Poole hospital, and from all patients who subsequently have been diagnosed with CRC. The **two datasets** consist of five main groups of variables, including:

- Demographic data: age, gender, date of assessment.
- Cancer data: TNM stage, histology, site, date of diagnose.
- Outcome of the investigations: none, cancer (3 types), adenoma, inflammation, vascular malformation, coeliac.
- Investigations: none is done, colonoscopy, gastroscopy (OGD).
- Blood test results: MCV, Hb, ferritin, T.satu.

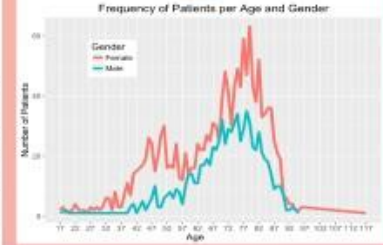
Methods: R programming language will be used for the statistical analysis. Then, two major branches of machine learning, namely classification and clustering, will be applied in this study.

A Box Plot of Blood Results per Gender and then per Age Interval



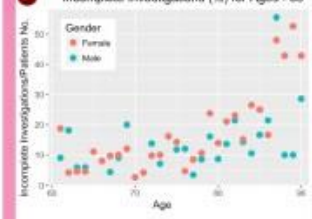
The Positive Outcomes of the Investigations

Frequency of Patients per Age and Gender



4. Initial Findings

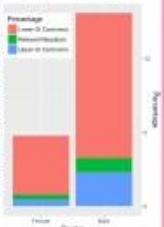
Incomplete Investigations (%) for Ages >60



During the whole period of the study, there was 323 non-completed investigation case due to patient preference or unsuitability. As can be shown from the graph above, and regardless the gender, the older patients are more likely to prefer not to complete the investigations or to be considered as unsuitable for the investigations.

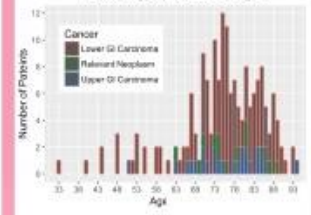
- There was 175 cancer cases, 41% of them were for female and 59% for male cases. As can be seen from the chart to the right, female are less likely to be diagnosed with any type of GI cancer than male patients. However, when diagnosed, female patients are more likely to have the lower GI cancer.

Bar Chart Of Cancer Types Percentages



- Lower GI carcinoma was the most frequent type of cancer in the study (137 case) and it can be diagnosed at younger ages as can be seen from the next graph. Upper GI carcinoma is the second most frequent type of cancer (25 case). Relevant neoplasm was the least frequent type and it starts at older ages.

Cancer Type per Patients' Ages



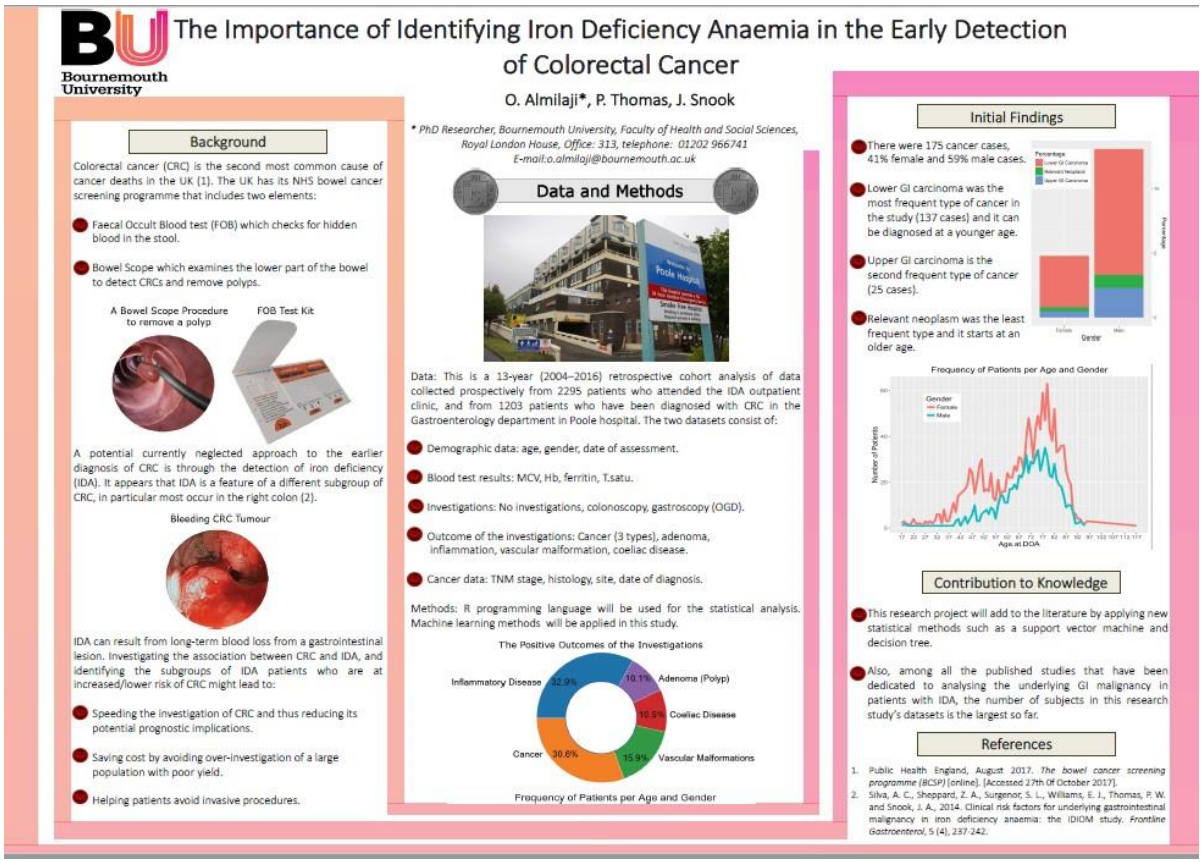
5. Contribution to Knowledge

- This research project would add to the literature by applying new statistical methods such as support vector machine and decision tree.
- Also, among all the published studies that are dedicated to analysing the underling GI malignancy in patients with IDA, the number of subjects in this research study's datasets is the largest so far.

References:
1. Public Health England, August 2017. The bowel cancer screening programme (BCSP), [Accessed 27th of October 2017].
2. Silva, A. C., Sheppard, J. A., Surgeon, S. L., Williams, E. I., Thomas, P. W. and Snook, J. A., 2014. Clinical risk factors for underlying gastrointestinal malignancy in iron deficiency anaemia: the IDEOM study. Frontline Gastroenterol, 5 (4), 237-242.

Appendix Ia2: BU PGR Live Exhibition. Dec 2018

Presented poster:



The Importance of Identifying Iron Deficiency Anaemia in the Early Detection of Colorectal Cancer

Bournemouth University

O. Almilaji*, P. Thomas, J. Snook

* PhD Researcher, Bournemouth University, Faculty of Health and Social Sciences, Royal London House, Office: 313, telephone: 01202 966741
E-mail:o.almilaji@bournemouth.ac.uk

Background

Colorectal cancer (CRC) is the second most common cause of cancer deaths in the UK (1). The UK has its NHS bowel cancer screening programme that includes two elements:

- Faecal Occult Blood test (FOB) which checks for hidden blood in the stool.
- Bowel Scope which examines the lower part of the bowel to detect CRCs and remove polyps.

A Bowel Scope Procedure to remove a polyp

FOB Test Kit

A potential currently neglected approach to the earlier diagnosis of CRC is through the detection of iron deficiency (IDA). It appears that IDA is a feature of a different subgroup of CRC, in particular most occur in the right colon (2).

Bleeding CRC Tumour

IDA can result from long-term blood loss from a gastrointestinal lesion. Investigating the association between CRC and IDA, and identifying the subgroups of IDA patients who are at increased/lower risk of CRC might lead to:

- Speeding the investigation of CRC and thus reducing its potential prognostic implications.
- Saving cost by avoiding over-investigation of a large population with poor yield.
- Helping patients avoid invasive procedures.

Data and Methods

Data: This is a 13-year (2004–2016) retrospective cohort analysis of data collected prospectively from 2295 patients who attended the IDA outpatient clinic, and from 1203 patients who have been diagnosed with CRC in the Gastroenterology department in Poole hospital. The two datasets consist of:

- Demographic data: age, gender, date of assessment.
- Blood test results: MCV, Hb, ferritin, T.satu.
- Investigations: No investigations, colonoscopy, gastroscopy (OGD).
- Outcome of the investigations: Cancer (3 types), adenoma, inflammation, vascular malformation, coeliac disease.
- Cancer data: TNM stage, histology, site, date of diagnosis.

Methods: R programming language will be used for the statistical analysis. Machine learning methods will be applied in this study.

The Positive Outcomes of the Investigations

Frequency of Patients per Age and Gender

Initial Findings

- There were 175 cancer cases, 41% female and 59% male cases.
- Lower GI carcinoma was the most frequent type of cancer in the study (137 cases) and it can be diagnosed at a younger age.
- Upper GI carcinoma is the second frequent type of cancer (25 cases).
- Relevant neoplasm was the least frequent type and it starts at an older age.

Frequency of Patients per Age and Gender

Contribution to Knowledge

- This research project will add to the literature by applying new statistical methods such as a support vector machine and decision tree.
- Also, among all the published studies that have been dedicated to analysing the underlying GI malignancy in patients with IDA, the number of subjects in this research study's datasets is the largest so far.

References

1. Public Health England, August 2017. *The bowel cancer screening programme (BCSP)* [online]. [Accessed 27th Of October 2017].
2. Silva, A. C., Sheppard, Z. A., Surgenor, S. L., Williams, E. J., Thomas, P. W. and Snook, J. A., 2014. Clinical risk factors for underlying gastrointestinal malignancy in iron deficiency anaemia: the IDIOM study. *Frontline Gastroenterol*, 5 (4), 237-242.

Appendix Ia3: British Society of Gastroenterology (BSG) Annual Meeting. Glasgow Jun 2019

Accepted abstract:

Open access published at: https://gut.bmj.com/content/68/Suppl_2/A192.1

Citation: Almilaji, O., Thomas, P., & Snook, J. 2019. PWE-042 Predicting the risk of gastrointestinal cancer in iron deficiency anaemia. *Gut* 68**:**A192.

PWE-042 Predicting the risk of gastro-intestinal cancer in iron deficiency anaemia

Introduction: Iron deficiency anaemia (IDA) is a common clinical presentation, and in a significant minority of cases (–0%) is the first indication of an underlying cancer in the gastro-intestinal (GI) tract. IDA is therefore considered an indication for fast-track endoscopic investigation, though the majority of cases will not actually have cancer. This study explores whether cancer risk in IDA can be predicted on the basis of simple and objective clinical variables.

Method: A study of the predictive value of sex, age, haemoglobin concentration (Hb), mean red cell volume (MCV) and iron studies for the risk of GI malignancy on subsequent investigation in adults with confirmed IDA attending a single IDA clinic. The study population comprised a training dataset (n = 2295) and a validation dataset (n = 602). The analysis was undertaken using logistic regression, and an App to predict the probability of GI cancer in IDA was developed as a clinical tool using R Shiny programming language.

Results: Using the training data, the best model showed that the risk of GI malignancy was strongly associated with sex (OR for males: 2.83, P<0.001) age, (OR: 1.05 for each added year, and Hb (OR: 0.975 for each g/l fall, P<0.001) – see figure 1 for the combined effects with their confidence intervals. GI cancer risk was less strongly associated with MCV (OR: 0.971 for each fl fall, P<0.05), with a complex relationship largely due to an

increased cancer risk in those with more severe anaemia, particularly in younger age-groups.

The model was tested on the validation data and produced similar results. It allowed stratification of 13% of the study population into a sub-group at high risk of cancer (arbitrarily defined as >15%), 28% into a sub-group at low risk (–%), and 16% into a sub-group at very low risk (<1%).

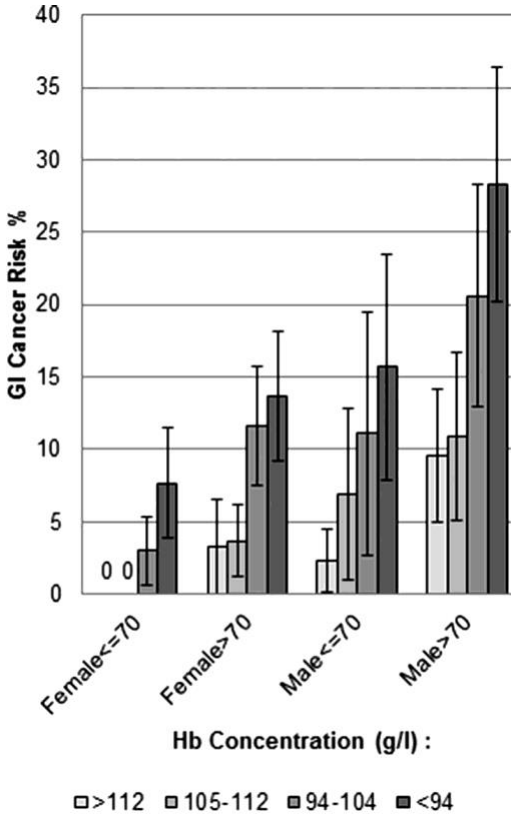


Figure 1

Presented poster:

Predicting the Risk of Gastro-Intestinal Cancer in Iron Deficiency Anaemia

Poole Hospital NHS Foundation Trust

Orouba Almilaji [1,2], Prof Peter Thomas [1], Dr Jonathon Snook [2]

[1]Bournemouth University, HSS, Clinical Research Unit
[2]Gastroenterology Department, Poole Hospital

Bournemouth University

Introduction

Iron deficiency anaemia (IDA) is a common clinical presentation, and in a significant minority of cases (8-10%) is the first indication of an underlying cancer in the gastro-intestinal (GI) tract.

IDA is therefore considered an indication for fast-track endoscopic investigation, though the majority of cases will not actually have cancer.

This study explores whether cancer risk in IDA can be predicted on the basis of simple and objective clinical variables.

Identifying that the patient has high risk of underlying GI malignancy, could warrant fast-track investigation. Similarly, identifying that the patient has risk low risk, means that his/her symptoms could reasonably be managed without the need for invasive investigation.

Aims

Deriving a score corresponding to the percentage probability of underlying GI malignancy in every IDA patient through a simple developed medical software application might lead to:

1. speeding the investigation of GI cancer. And thus reducing its potential prognostic implications.
2. saving cost by avoiding over-investigation of a large population with poor yield.
3. helping patients avoid invasive procedures.

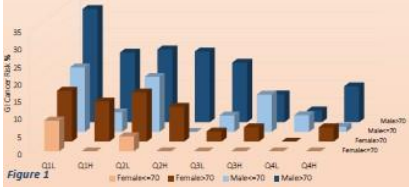


Figure 1

Figure 2

Methods

A study of the predictive value of sex, age, haemoglobin concentration (Hb), mean red cell volume (MCV) and iron studies for the risk of GI malignancy on subsequent investigation in adults with confirmed IDA attending a single IDA clinic.

The study population comprised a training dataset (n = 2295) and a validation dataset (n = 602).

The analysis was undertaken using logistic regression, and an App (Predict Logic-IDA) to predict the probability of GI cancer in IDA was developed tool using R Shiny programming language based on merging the two datasets to build the final model for this tool.

The tool is classified as medical device class one and will be used by health professional in order to aid their decisions with regard to which clinical pathway is to be followed.

Results

Using the training data, the best model showed that the risk of GI malignancy was strongly associated with sex (OR for males: 2.83, P<0.001) age, (OR: 1.05 for each added year, and Hb (OR: 0.975 for each g/l fall, P<0.001) – see figure 3 for the combined effects with their confidence intervals.

GI cancer risk was less strongly associated with MCV (OR: 0.971 for each fl fall, P<0.05), with a complex relationship largely due to an increased cancer risk in those with more severe anaemia, particularly in younger age-groups – **see figure 1.**

The model was tested on the validation data and produced similar results. It allowed stratification of 13% of the study population into a sub-group at high risk of cancer (arbitrarily defined as >15%), 28% into a sub-group at low risk (1-5%), and 16% into a sub-group at very low risk (<1%).

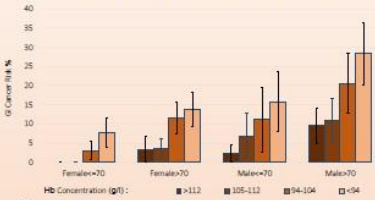


Figure 3

Conclusions

This study confirms that a simple clinical scoring system can effectively stratify patients with IDA according to GI cancer risk, allowing stretched investigational resources to be targeted at the high-risk group, whilst perhaps avoiding invasive investigation altogether in those predicted to be at extremely low risk.

The App developed has the potential to provide a quick estimate of GI cancer risk in clinical settings, and so facilitate patient counselling.

Appendix Ia4: NCRI Virtual Showcase. Nov 2020

Accepted abstract:

Open access published at: https://abstracts.ncri.org.uk/abstract/external-validation-of-the-idiom-score-for-predicting-the-risk-of-gastro-intestinal-malignancy-in-iron-deficiency-anaemia/

Citation: Almilaji, O., Webb, G., Chapman, T. P., Williams, E. J., Shine, B. S. F., Ellis, A. J., Docherty, S., & Snook, J. Internal and External validation of the IDIOM score for predicting the risk of gastrointestinal malignancy in iron deficiency anaemia. *National Cancer Research Institute (NCRI) Virtual Showcase.* 2 -3 Nov 2020

External and Internal validation of the IDIOM score for predicting the risk of gastro-intestinal malignancy in iron deficiency anaemia

Background: Gastrointestinal (GI) malignancy is a common finding in iron deficiency anaemia (IDA), with a prevalence of about 8%. Using two large datasets from Dorset, we have previously reported and internally validated a model for predicting the risk of GI malignancy in IDA – the IDIOM score. This is based on four independent and objective clinical parameters - age, sex, mean corpuscular volume (MCV), and haemoglobin concentration (Hb). This study aims to assess the performance of the predictive model applied to an unrelated external validation dataset.

Method: The external validation dataset was derived from a different population (in Oxford), collected under different circumstances (from fast-track referrals), and comprised a total of 1118 patients with confirmed IDA. The data were anonymised prior to analysis. The logistic regression model based on the training data was used to predict the GI malignancy risk in this new dataset. Due to the imbalance between the “positive” and “negative” GI malignancy numbers, geometric mean (G mean), and negative predictive value were used to assess the performance of the model.

Results: The characteristics of the external validation dataset differed from those of the training dataset, with lower mean Hb in particular. Using the

regression model to calculate predicted GI malignancy risk, a threshold risk of 7.43% maximised the G mean in the training dataset (69%) and gave a comparable value in the external validation dataset (61%). At this threshold, sensitivity and specificity were 76% and 63% respectively in the training dataset and 84% and 44% in the external validation dataset. A predicted risk threshold of 1.5% was the largest value to give a negative predictive value of 100% in the training dataset and gave an identical result in the external validation dataset.

Conclusion: This external validation exercise has demonstrated that the model underlying the IDIOM score is robust in predicting the risk of underlying GI malignancy in a large IDA dataset collected in a different clinical setting.

Impact statement: Ultimately, validating the model would help to using it to rationalise the use of investigational resources in IDA, by fast-tracking high-risk patients and, with appropriate safeguards, avoiding invasive investigation altogether in those at ultra-low predicted risk.

Presented poster:

External and Internal Validation of the IDIOM Score for Predicting the Risk of Gastro-intestinal Malignancy in Iron Deficiency Anaemia

Orouba Almilaji [1,3], Gwilym Webb[2], Thomas P Chapman[2], Elizabeth J Williams[3], Brian SF Shine[4], Antony J Ellis[2], Sharon Docherty[1], Jonathon Snook[3]

[1] *Medical Science and Public Health Department, HSS, Bournemouth University*
[2] *Translational Gastroenterology Unit, NIHR Oxford Biomedical Research Centre, University of Oxford*
[3] *Gastroenterology Department, Poole Hospital*
[4] *Nuffield Department of Clinical Laboratory Sciences, University of Oxford*

Bournemouth University

NHS Poole Hospital NHS Foundation Trust

University of Oxford

Background

Gastrointestinal (GI) malignancy is a common finding in iron deficiency anaemia (IDA), with a prevalence of about 8%. Using One large dataset from Dorset (training dataset), we have previously reported a model for predicting the risk of GI malignancy in IDA – the IDIOM score. This is based on four independent and objective clinical parameters - age, sex, mean corpuscular volume (MCV), and haemoglobin concentration (Hb). This study aims to validate the performance of the predictive model applied to an internal similar dataset, and to an unrelated external dataset.

Methods

The two datasets comprised of patients with confirmed IDA. The internal dataset was derived from a similar population (in Dorset), and collected during a different time. The external dataset was derived from a different population (in Oxford), and collected under different circumstances (from fast-track referrals). The data were anonymised prior to analysis. The logistic regression model based on the training data was used to predict the GI malignancy risk in the internal and external validation datasets. Due to the imbalance between the "positive" and "negative" GI malignancy numbers, geometric mean (G mean) of sensitivity and specificity, and negative predictive value were used to assess the performance of the model.

NCRI National Cancer Research Institute

Partners in cancer research

Table 1

		Training dataset (Dorset)	Internal validation dataset (Dorset)	External validation dataset (Oxford)
Number		1879	511	1118
GI cancer	positive - n (%)	157 (8.4%)	43 (8.4%)	86 (7.7%)
Sex ratio	M/F	0.56	0.59	0.67
Age (years)	median (Q1, Q3)	71 (59, 78)	71 (61, 79)	74 (65, 81)
Hb (g/l)	median (Q1, Q3)	104 (93, 112)	107 (95, 118)	91 (79, 101)
MCV (fl)	median (Q1, Q3)	80 (74, 86)	83 (77, 89)	81 (75, 87)

Results

The characteristics of the internal validation dataset is similar to those of the training dataset. The characteristics of the external validation dataset differed from those of the training dataset (**Table 1**), with lower mean (and median) Hb in particular. Using the regression model to calculate predicted GI malignancy risk, a threshold risk of 7.43% maximised the G mean in the training dataset (69%) and gave comparable values in the internal dataset (73%), and in the external dataset (61%). At this threshold, sensitivity and specificity were 76% and 63% respectively in the training dataset, 79% and 68% in the internal validation dataset, and 84% and 44% in the external validation dataset. A predicted risk threshold of 1.5% was the largest value to give a negative predictive value of 100% in the training dataset and gave an identical result in the internal and external validation datasets.

Conclusions

This validation exercise has demonstrated that the model underlying the IDIOM score is robust in predicting the risk of underlying GI malignancy in large IDA datasets collected in a different time frame, and in a different clinical setting.

conference.ncri.org.uk #NCRIVirtual

Appendix Ia5: CRUK Early Detection of Cancer Conference. Oct 2020, Online

Accepted abstract (non-open access**):**

Citation: Almilaji, O., Webb, G., Chapman, T.P., Williams, E.J., Shine, B.S.F., Ellis, A.J., Docherty, S. and Snook, J., 2020. External validation of the IDIOM score for predicting the risk of gastro-intestinal malignancy in iron deficiency anaemia. In: CRUK Early Detection of Cancer 6-8 October 2020 online.

Presented poster:

This image has been redacted as the publication is not available without registration with the conference website.

Appendix Ia6: American College of Gastroenterology (ACG) Annual Scientific Meeting. Oct 2021

Accepted abstract:

Open access published at:

https://journals.lww.com/ajg/Fulltext/2021/10001/S1313_Modelling_the_Episodes_of_Care_for_IDA.1317.aspx

Citation: Almilaji, O., S1313 Modelling the Episodes of Care for IDA Patients in a Secondary Care Centre Using Continuous-Time Multistate Markov Chain, The American Journal of Gastroenterology: October 2021 - Volume 116 - Issue - p S605 doi: 10.14309/01.ajg.0000778784.81979.c6.

S1313 Modelling the Episodes of Care for IDA Patients in a Secondary Care Centre Using Continuous-Time Multistate Markov Chain

Introduction: Despite the high prevalence of gastro-intestinal (GI) cancer in iron deficiency anaemia (IDA), some IDA patients do not complete all the necessary GI investigations at the initial referral due to informed patient preference, concurrent illness, or major co-morbidity including frailty. As a result, existing cancers are diagnosed at a later referral with worse prognosis. The potential to detect GI cancer early depends on minimizing the delay time spent between the two consecutive referrals, where a patient did not complete investigations at the first referral, but at the second is diagnosed with positive GI cancer. This retrospective longitudinal study aims to highlight the proper methods to model these referrals.

Methods: Using anonymised data of 168 episodes of care for IDA patients at an IDA clinic that was established under the supervision of the Gastroenterology Department at General Hospital, continuous-time multi-state Markov chain is employed to determine the transition rates between three observed states for IDA patients at the IDA clinic; “incomplete investigations”, “negative GI cancer”, and “positive GI cancer” and to estimate the delay time.

Results: Once in the state of incomplete investigations, an estimated mean delay time of 3.1 years (95% CI: 1.2, 5) is spent before being diagnosed with positive GI cancer. The probability that the “positive GI diagnosis” is next after the state of “incomplete investigation” is 0.17 (95% CI: 0.03, 0.54) compared with 11% (95% CI: 0.02,0.39) when it is followed the state of negative GI cancer. The survival rate of IDA patients with negative GI diagnosis is always higher than those with incomplete investigations (Figure 1). Finally, being diagnosed with positive GI cancer is always preceded by the prediction of being considered “very high risk” at the earlier visit.

Conclusion: A baseline model was developed to represent episodes of care for IDA patients at a secondary care centre. The suggested methodology in this study can be used in the future to help policy makers establishing what is the maximum delay time, a confirmed IDA patient, should not be allowed to stay in before investigated for GI cancer, and what are the measures that could be put in place to reduce this time.

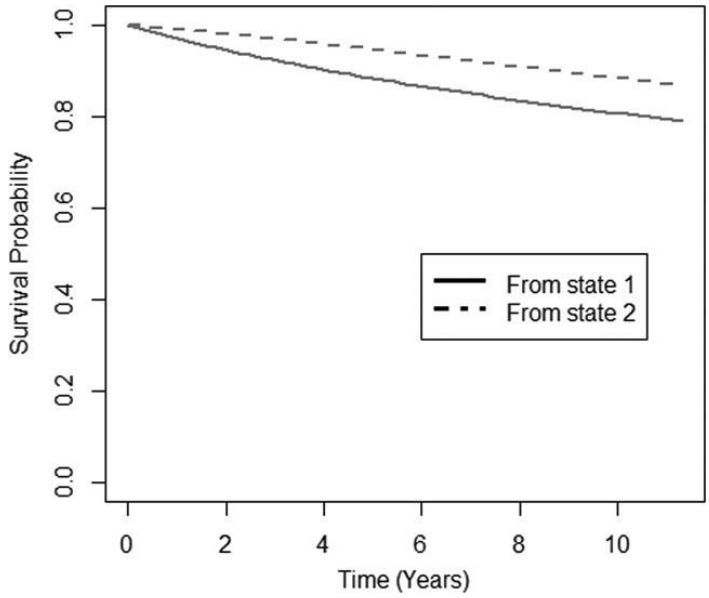


Figure 1. Survival plot. Survival is defined as not being in the state of “positive GI cancer”. State 1 means being observed with incomplete investigations. state 2 means being observed with negative GI cancer.

Presented poster:

Modelling The Episodes of Care for IDA Patients in A Secondary-Care Centre Using Continuous-Time Multistate Markov Chain

Orouba Almilaji [1,2]

[1] Gastroenterology Unit, University Hospitals Dorset NHS Foundation Trust, Poole, UK

[2] Department of Medical Science and Public Health, Bournemouth University, Bournemouth, UK

E-mail: oalmilaji@bournemouth.ac.uk

INTRODUCTION: Despite the high prevalence of gastro-intestinal (GI) cancer in iron deficiency anaemia (IDA), some IDA patients do not complete all the necessary GI investigations at the initial referral (figure 1). As a result, existing cancers are diagnosed at a later referral with worse prognosis. The potential to detect GI cancer early depends on minimizing the delay time spent between the two consecutive referrals, where a patient did not complete investigations at the first referral, but at the second is diagnosed with positive GI cancer. This retrospective longitudinal study aims to highlight the proper methods to model these referrals.

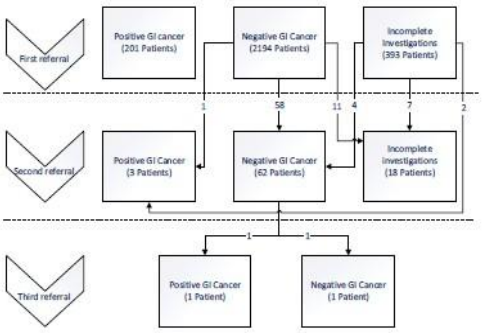


Figure 1 Flow chart of patient states at the IDA clinic during the study period 2004-2018

METHODS: Using anonymised data of 168 episodes of care for IDA patients at an IDA clinic in secondary care setting, continuous-time multi-state Markov chain is employed to determine the transition rates between three observed states for IDA patients at the IDA clinic; "incomplete investigations", "negative GI cancer", and "positive GI cancer" and to estimate the delay time.

RESULTS: Once in the state of incomplete investigations, an estimated mean delay time of 3.1 years (95% CI: 1.2, 5) is spent before being diagnosed with positive GI cancer. The probability that the "positive GI diagnosis" is next after the state of "incomplete investigation" is 17% compared with 11% when it is followed the state of negative GI cancer. Defining the survival as the event of not being in the state of "positive GI cancer", the survival rate of IDA patients with negative GI cancer is always higher than those with incomplete investigations (figure 2). Finally, being diagnosed with positive GI cancer according to the IDIOM score[1,2] is always preceded by the prediction of being considered "very high risk" at the earlier visit.

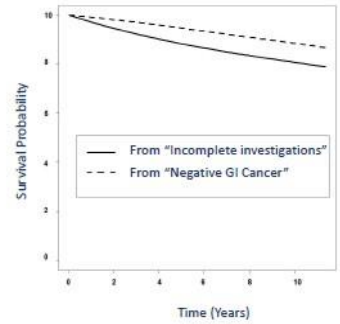


Figure 2 Survival plot. Survival is defined as not being in the state of "positive GI cancer".

CONCLUSION: A baseline model was proposed to represent episodes of care for IDA patients at a secondary care centre. Preliminary results highlight the importance of completing the GI investigations especially in IDA patients who are at high risk of GI cancer and fit enough to do the investigations.

Ref 1: Almilaji O, Snook JA, Thomas P. The IDIOM APP, 2020. Available: https://www. predict-gi-risk-in-ida.com [Accessed 4 Oct 2021].
Ref 2: 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 Gastroenterol* 2020;7:e000403.

Appendix Ia7: Virtual NCRI Festival: Making Cancer Research Better Together. Nov 2021

Accepted abstract:

Open access published at:

https://abstracts.ncri.org.uk/abstract/modelling-the-episodes-of-care-for-iron-deficiency-anaemia-patients-in-a-secondary-care-centre-using-continuous-time-multi-state-markov-chain/

Citation: Almilaji, O., Docherty, S., & Snook, J. Modelling the episodes of care for IDA patients in a secondary-care centre using continuous-time multistate Markov chain. (ID: 3556). *Virtual NCRI Festival: Making cancer research better together*. 8-12 Nov 2021.

Modelling the episodes of care for iron deficiency anaemia patients in a secondary-care centre using continuous-time multi-state Markov chain

Background: Due to informed patient preference, concurrent illness, or major co-morbidity, some IDA patients do not complete all the necessary GI investigations at the initial referral . As a result, existing cancers are diagnosed at a later referral with worse prognosis. The potential to detect GI cancer early depends on minimizing the delay time spent between the two consecutive referrals. This study aims to highlight the proper methods to model these referrals.

Method: Using anonymised data of 168 episodes of care for IDA patients, continuous-time multi-state Markov chain is employed to determine the transition rates between three observed states for IDA patients; “incomplete investigations”, “negative GI cancer”, and “positive GI cancer”; and to estimate the delay time.

Results: Once in the state of incomplete investigations, an estimated mean delay time of 3.1 years (95% CI: 1.2, 5) is spent before being diagnosed with positive GI cancer. The probability that the “positive GI diagnosis” is next after the state of “incomplete investigation” is 17% compared with 11% when it is

followed the state of negative GI cancer. Finally, being diagnosed with positive GI cancer is always preceded by the prediction[1, 2, 3] of being considered "very high risk" at the earlier visit.

Conclusion: A baseline model was developed to represent episodes of care for IDA patients at a secondary care centre.

Impact statement: The suggested methodology can be used in the future to help policy makers establishing what is the maximum delay time, a confirmed IDA patient, should not be allowed to stay in before investigated for GI cancer, and what are the measures that could be put in place to reduce this time.

References:

1: Almilaji O, Snook J, Thomas P. 2020. The IDIOM App. https://www.predict-gi-risk-in-ida.com

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

3: Almilaji O, Parry SD, Docherty S, Snook J. 2021. Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia. *Sci Rep*. 11(1):13055. doi: 10.1038/s41598-021-92623-z

Presented poster:

Modelling The Episodes of Care For Iron Deficiency Anaemia Patients in A Secondary-care Centre Using Continuous-time Multi-state Markov Chain

Orouba Almilaji[12], Sharon Docherty[2], Jonathon Snook[1]

[1] Gastroenterology Unit, University Hospitals Dorset NHS Foundation Trust, Poole, UK

[2] Department of Medical Science and Public Health, Bournemouth University, Bournemouth, UK

E-mail: oalmilaji@bournemouth.ac.uk

Background

Despite the high prevalence of gastro-intestinal (GI) cancer in iron deficiency anaemia (IDA), some IDA patients do not complete all the necessary GI investigations at the initial referral due to informed patient preference, concurrent illness, or major co-morbidity including frailty. As a result, existing cancers are diagnosed at a later referral with worse prognosis. The potential to detect GI cancer early depends on minimizing the delay time spent between the two consecutive referrals, where a patient did not complete investigations at the first referral, but at the second is diagnosed with positive GI cancer. This retrospective longitudinal study aims to highlight the proper methods to model these referrals.

Method

Using anonymised data of 168 episodes of care for IDA patients at an IDA clinic that was established under the supervision of the Gastroenterology Department at General Hospital, continuous-time multi-state Markov chain is employed to determine the transition rates between three observed states for IDA patients at the IDA clinic; "incomplete investigations", "negative GI cancer", and "positive GI cancer" and to estimate the delay time.

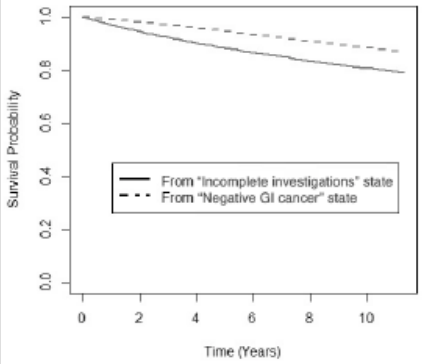


Figure 1 Survival plot. Survival is defined as not being in the state of "positive GI cancer".

Results

Once in the state of incomplete investigations, an estimated mean delay time of 3.1 years (95% CI: 1.2, 5) is spent before being diagnosed with positive GI cancer. The probability that the "positive GI diagnosis" is next after the state of "incomplete investigation" is 17% compared with 11% when it is followed the state of negative GI cancer. The survival rate of IDA patients with negative GI diagnosis is always higher than those with incomplete investigations (figure 1). Finally, being diagnosed with positive GI cancer is always preceded by the prediction by the IDIOM score[1,2] of being considered "very high risk" at the earlier visit.

Conclusions

A baseline model was proposed to represent episodes of care for IDA patients at a secondary care centre. The suggested methodology in this small-size study can be used in the future to help policy makers establishing what is the maximum delay time, a confirmed IDA patient, should not be allowed to stay in before investigated for GI cancer, and what are the measures that could be put in place to reduce this time using appropriate study sample size.

References:

1: Almilaji O, Snook JA, Thomas P. The IDIOM APP, 2020. Available: https://www. predict-gi-risk-in-ida.com [Accessed 4 Oct 2021].

2: 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 Gastroenterol* 2020;7:e000403.

Appendix Ia8: British Society of Gastroenterology (BSG) Campus. Jan 2021, Online

Accepted abstract:

Open access published at:

https://gut.bmj.com/content/70/Suppl_1/A37.2

Citation: Almilaji, O., Engen, V., Snook, J., & Thomas, P., 2021. The development of a web-based application to predict the risk of GI cancer in IDA. https://youtu.be/DRjenumKdhY. *Gut* 70:A37-A38.

O67 The development of a web-based application to predict the risk of GI cancer in IDA

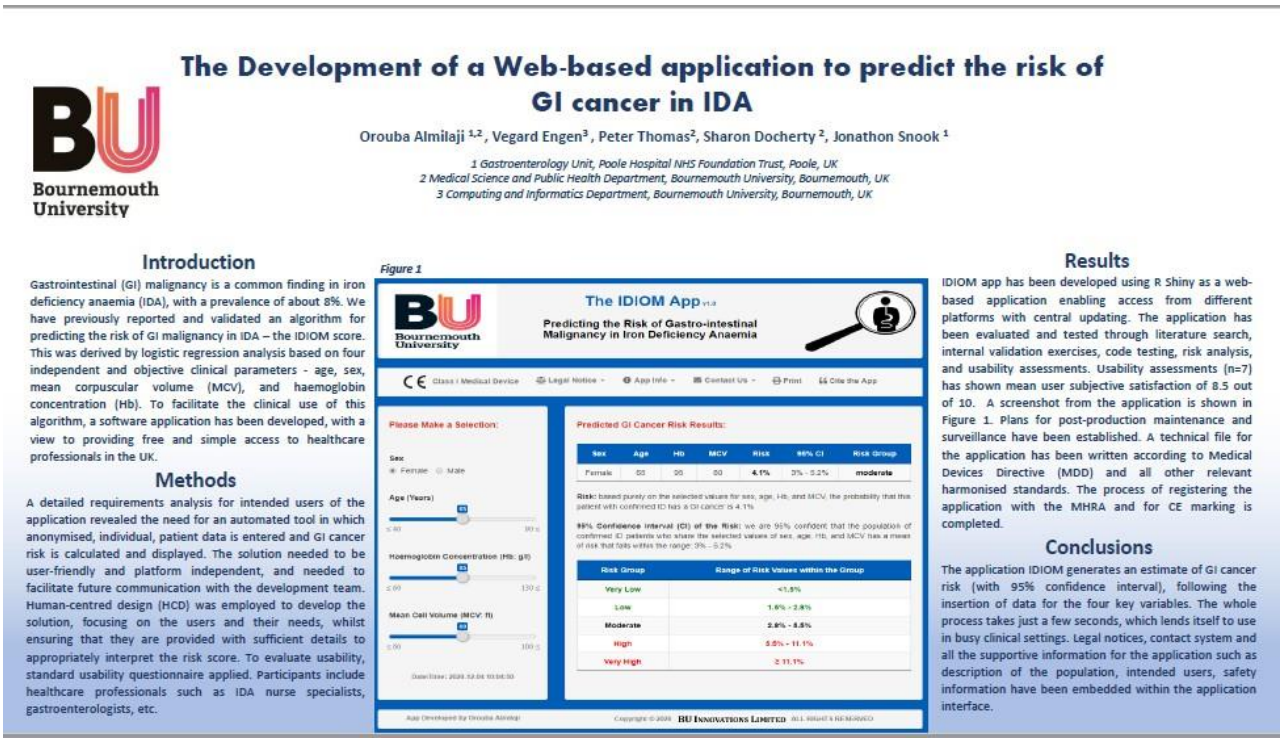
Introduction: Gastrointestinal (GI) malignancy is a common finding in iron deficiency anaemia (IDA), with a prevalence of about 8%. We have previously reported and validated an algorithm for predicting the risk of GI malignancy in IDA – the IDIOM score. This was derived by logistic regression analysis based on four independent and objective clinical parameters - age, sex, mean corpuscular volume (MCV), and haemoglobin concentration (Hb). To facilitate the clinical use of this algorithm, a software application has been developed, with a view to providing free and simple access to healthcare professionals in the UK.

Methods: A detailed requirements analysis for intended users of the application revealed the need for an automated tool in which anonymised, individual, patient data is entered and GI cancer risk is calculated and displayed. The solution needed to be user-friendly and platform independent, and needed to facilitate future communication with the development team. Human-centred design (HCD) was employed to develop the solution, focusing on the users and their needs, whilst ensuring that they are provided with sufficient details to appropriately interpret the risk score. To evaluate usability, standard usability questionnaire applied. Participants include healthcare professionals such as IDA nurse specialists, gastroenterologists, etc.

Results: *Predict GI Cancer in IDA* has been developed using R Shiny as a web-based application enabling access from different platforms with central updating. The application has been evaluated and tested through literature search, internal validation exercises, code testing, risk analysis, and usability assessments. Usability assessments (n=7) has shown mean user subjective satisfaction of 8.5 out of 10. Plans for post-production maintenance and surveillance have been established. A technical file for the application has been written according to Medical Devices Directive (MDD) and all other relevant harmonised standards. The process of registering the application with the MHRA and for CE marking is underway.

Conclusions: The application *Predict GI Cancer in IDA* generates an estimate of GI cancer risk (with 95% confidence interval), following the insertion of data for the four key variables. The whole process takes just a few seconds, which lends itself to use in busy clinical settings. Legal notices, contact system and all the supportive information for the application such as description of the population, intended users, safety information have been embedded within the application interface.

Presented poster:



Appendix Ia9: British Society of Gastroenterology (BSG) Campus. Jan 2021, Online

Accepted abstract:

Open access published at:

https://gut.bmj.com/content/70/Suppl_1/A190.1

Citation: Almilaji, O., Parry, S., Thomas, P. and Snook, J., 2021. Downstaging of right-sided colorectal cancer diagnosed through iron deficiency anaemia. https://youtu.be/R7_llJjRBEo. *Gut*;70:A190.

P288 Downstaging of right-sided colorectal cancer diagnosed through iron deficiency anaemia

Introduction: Previous studies have suggested that iron deficiency anaemia (IDA) is an indicator of poor prognosis in colorectal cancer (CRC), but this may be due to confounding – IDA is much commoner in right-sided CRC, which tends to late presentation and therefore a worse prognosis. This study aims to determine the effect of diagnosing CRC through the detection of IDA on tumour stage - a surrogate marker of prognosis in CRC - whilst controlling for tumour side.

Methods: A total of 1154 cases of CRC with adequate clinical information were identified from the MDT records of a single general hospital for 2010–2016. Histological confirmation of adenocarcinoma was available in 90%. Each case was staged on the basis of the available radiological and surgical evidence, and the route of presentation identified. Because tumour side and presentation are surrogate markers of prognosis in CRC, these variables were merged to create a new variable to reflect CRC prognosis, and analysed using binary logistic regression models.

Results: A summary of the basic patient data is shown in table 1.

Table 1

	IDA	Screening	Symptomatic	Overall
Number	171	213	770	1154
Sex ratio – M/F	1.1	1.5	1.3	1.3
Age (years) – mean (sd)	77 (± 11)	68 (± 6)	73 (± 13)	72 (± 12)
Hb (g/l) – mean (sd)	88 (± 17)	133 (± 19)	122 (± 23)	119 (± 25)
Early stage (I or II) – n (%)	89 (52.0%)	127 (59.6%)	304 (39.5%)	520 (45.1%)
Right-sided – n (%)	141 (82.5%)	71 (33.3%)	243 (31.6%)	455 (39.4%)

As anticipated, most cases presenting with IDA proved to have right-sided tumours, whilst the majority of cases diagnosed through screening were left-sided. As expected, left-sided tumours diagnosed through screening (mostly in the national bowel cancer screening programme) were significantly down-staged in comparison to those presenting with symptomatic disease – with an odds ratio for early stage disease of 2.09 (95% CI 1.4 - 3.1, P<0.001).

The key finding in this study is that right-sided tumours diagnosed following the detection of IDA also appear to be down-staged compared to those presenting with symptomatic disease – with an odds ratio for early stage disease of 2.52 (95% CI 1.6 - 3.8, P<0.0001).

Conclusion: The findings suggest a prognostic benefit to diagnosing right-sided CRC through the detection of IDA, with a benefit comparable to that of the screening programme for left-sided CRC. This strengthens the case for a systematic approach to blood count monitoring in the population at-risk of CRC.

Presented poster:

Downstaging of Right-sided Colorectal Cancer Diagnosed Through Iron Deficiency Anaemia

Orouba Almilaji [1,2], Sally D Parry [1], Peter Thomas [2], Sharon Docherty [2], Jonathon Snook [1]

1 Gastroenterology Unit, Poole Hospital NHS Foundation Trust, Poole, UK
2 Medical Science and Public Health Department, Bournemouth University, Bournemouth, UK

Bournemouth University

NHS Poole Hospital NHS Foundation Trust

Introduction

Previous studies have suggested that iron deficiency anaemia (IDA) is an indicator of poor prognosis in colorectal cancer (CRC), but this may be due to confounding – IDA is much commoner in right-sided CRC, which tends to late presentation and therefore a worse prognosis. This study aims to determine the effect of diagnosing CRC through the detection of IDA on tumour stage - a surrogate marker of prognosis in CRC - whilst controlling for tumour side.

Methods

A total of 1154 cases of CRC with adequate clinical information were identified from the MDT records of a single general hospital for 2010-2016. Histological confirmation of adenocarcinoma was available in 90%. Each case was staged on the basis of the available radiological and surgical evidence, and the route of presentation identified. Because tumour side and presentation are surrogate markers of prognosis in CRC, these variables were merged to create a new variable to reflect CRC prognosis, and analysed using binary logistic regression models.

Results

A summary of the basic patient data is shown in Table 1. As anticipated, most cases presenting with IDA proved to have right-sided tumours, whilst the majority of cases diagnosed through screening were left-sided.

Table1

	IDA	Screening	Symptomatic
Number	171	213	770
Sex ratio – M/F	1.1	1.5	1.3
Age (years) – mean (sd)	77 (± 11)	68 (± 6)	73 (± 13)
Hb (g/l) – mean (sd)	88 (± 17)	133 (± 19)	122 (± 23)
Early stage (I or II) – n (%)	89 (52.0%)	127 (59.6%)	304 (39.5%)
Right-sided – n (%)	141 (82.5%)	71 (33.3%)	243 (31.6%)

As expected, left-sided tumours diagnosed through screening (mostly in the national bowel cancer screening programme) were significantly down-staged in comparison to those presenting with symptomatic disease – with an odds ratio for early stage disease of 2.09 (95% CI 1.4 - 3.1, P<0.001).
The key finding in this study is that right-sided tumours diagnosed following the detection of IDA also appear to be down-staged compared to those presenting with symptomatic disease – with an odds ratio for early stage disease of 2.52 (95% CI 1.6 - 3.8, P<0.0001).

Conclusions

The findings suggest a prognostic benefit to diagnosing right-sided CRC through the detection of IDA, with a benefit comparable to that of the screening programme for left-sided CRC. This strengthens the case for a systematic approach to blood count monitoring in the population at-risk of CRC.

Appendix Ib: Declaration of Conformity Assessment

4. Declaration of Conformity

Manufacture Name:	BU Innovations Limited (BUI)
Manufacture address:	BU Innovations Limited, 2nd Floor Melbury House, 1-3 Oxford Road, Bournemouth BH8 8ES.
E-mail:	IDIOM.research.bucru@bournemouth.ac.uk IDIOM.maintenance.bucru@bournemouth.ac.uk
Phone Number:	01202 966741
Medical Device Name (Common name):	IDIOM (Iron Deficiency as an Indicator of Malignancy)
Device Description & Function Designation:	Predicting the risk of Gastro-Intestinal Cancer in Iron Deficiency Anaemia Patients through a Web-based Application.
Accessories:	None
Version number:	1.0
GMDN Code:	61087
Device Classification:	Class I, Active Stand-alone, Clinical decision-support (CDS) software, none-sterile, none-measuring, none-reusable surgical instrument.
Conformity Assessment Route:	Annex VII: Prepare technical documentation to support Declaration of Conformity

BU Innovations Limited hereby declares that the medical device specified above meets the provisions of the Council Directive; MDD No. 93/42/EEC of 14 June 1993 concerning medical devices as amended by Directive 2007/47/EC.
All supporting documentation is retained at the premises of BU Innovations Limited.

BU Innovations Limited Date of issue: 1st December 2020

Authorised representative name (Print)
Tim McIntyre-Bhatty:

Signature:

Appendix II: The Data cleaning

Appendix IIa: Cleaning Training Dataset (IDA)

- **Correction in year 2017**

1. Delete one header from the two headers namely the first one which includes: demographic data, dates, Initial Haematology, Investigation, and outcome.
2. Round the age to a whole number.
3. Change the birth year of case 2146 from 2016 to 1943.
4. Build three variables out of the two variables of Date of birth and Date of assessment. The new variables are in terms of (day, month, year).
5. Correct the levels of the gender variable from four levels (“f”, “f “, “F”,“m”) to two levels (Female, Male).
6. Change the values: "x", "X", "" in all blood test results variables to NA.
7. Change the Ferritin value from 0 to 8 in case 897.
8. Change the levels of OGD from 0, 1 to 1, 2. 0 showed as null values in R (228 null variables) and OGD levels are appeared as only one level.
9. Changed the values: "x", "" in Colonoscopy variable to 0, cases 355, and 2128.
10. Change the levels of Colonoscopy from 0, 1 to 1, 2 as 0 showed as null values in R (327 null variables) and Colonoscopy levels are appeared as only one level.
11. Change the NA in GI Investigation incomplete variable to 0 which means no-investigations, then change the levels from 0, 1, 2 to 1, 2, 3
12. Change the "x" in Cancer, Adenoma, Inflam, and Coeliac variables to NA, then change the levels from 0, 1, 2, 3 to 1, 2, 3, 4 according to the exiting levels, in Coeliac only two levels are existed: 0,1 to 1, 2. In Adenoma and Inflam only three levels existed, so 0, 1, 2 changed to 1, 2, 3.
13. Ordering the dataset according to the assessment date and change the values of the months from 1:12 to the names of the months.

14. Correct the day of assessment in observation 1387 from “ 26” to 26.

- **Correction in January 2018**

Inspecting the complete/incomplete GI investigation values has bought some observations into question, especially that, missing data for the outcome variables (321) is not equal to the 322 of incomplete assessments due to patient preference or unsuitability. There was need to check patient records for:

Cases (957, 1867, 1920) in which there are incomplete investigations due to patient preference or unsuitability, yet outcomes have values.

After checking with Dr. Snook, the following table has been given to correct the values:

Also, three observations (80, 1716 and 2207,) that I have asked to be checked have been confirmed to be correct. Cancers in observations 80, 1716 and 2207, indeed, have not been discovered by colonoscopy or OGD. Cancer in observation 80 discovered by laparotomy. Cancer in 2207 has been discovered by CT scan and confirmed by Ascitic cytology test. Cancer in 1716 has been discovered by CT scan and confirmed by Pleural effusion cytology test.

- **Correction in March 2018**

Change the MCV values in:

1813: from 24.6 to 78 6

317: from 26.3 to 81.8

1443: from 24 to 74.3

Duplicates have been found in the dataset as follows:

Repeated entries in which a patient has been seen at the clinic in the same day: 234,2018,544,1387,1385.

Repeated entries in which a patient has been seen at the clinic in the same year:

1244,821,487,793,1066,2125,568,1087,2087,1386,1860,2120,837,1788,1342,1531,1541,1650,1424,2138,743,1255,1872,949,2101,1487,2000,946,1812,1404

- **Correction in May 2018**

1. Change the data of birth from 21/4/32 to 21/4/52 for A2018.

2. Change the gender from female to male for a954, a1501, a532.

3. Delete case234 (no borrowing of values from it to substitute in the matched record).

4. Delete case1244 (no borrowing of values from it to substitute in the matched record).

5. Delete case 544 (substitute ferritin value of 8 in the case of 538 with the value. 7 from the deleted matched case of 544, cases recorded in the same day).

6. Delete case 2138 (no borrowing of values from it to substitute in the matched record).

7. Delete case 1385 (no borrowing of values from it to substitute in the matched record).

8. Delete case 1379 because it belongs to the same person who has been seen in the same day more than once).

9. Change case 2218 birthdate from 1/1/1900 to 19/5/1938.

10. Change case 1813 MCV from 24.6 to 78.

11. Change case 1443 MCV from 24 to 74.3.

12. Change case 317 MCV from 26.3 to 81.8.

Appendix IIb: Cleaning the Internal Validation Dataset

- **Correction on October 2018**

1. Taking out the first header (demographic data, Initial Haematology, Investigations, and outcomes).

2. Change the levels of Gender variable:

m to Male, M to Male

f to Female, F to Female

3. Round the age to whole number.

4. Change the values of MCV (less than 50) for 298, 283, and 377 (human entry error: the nurses entered the values of MCH instead of MCV).

5. Checked the date of birth dates for some young female demographic with very low blood indicators. (all correct).

6. Check the gender of 407, it was right and did not change: male.

7. The missing coeliac value of 553 which is 0.

8. Check the results of investigation for 112 since OGD =1 (was and still correct no need to change).

9. Check the odd values for T.satu in 483,414,529. All correct. T.satu responds quickly to iron treatment before ferritin and this is why we might have this 39 or 17.

10. Delete case v40 from the testing data as it is the same record for patient A2139 (identical) in the training dataset.

Appendix IIc: Cleaning CRC Dataset

- Initial size: 1258 rows * 11 variables

Variables: Study.No, Date.of.Birth, Gender, Date.of.Diagnosis, Diagnosis, Synch.tumours, Histology, Presentation, Haemoglobin, Final.Staging, TNM.Stage.

55 entries were taken out due to other neoplastic diagnoses and diagnosis of CRC and non-incident presentation made at another hospital. This left 1203 rows * 11 variables

- Duplicates: 14 records (7 patients):

Study IDS: "B451" , "B1183" , "B141" , "B421" , "B252" , "B749" , "B158" , "B1164" , "B426", "B901" , "B405" , "B623" , "B236" , "B625"

After keeping the earliest record per patient and removing the second records: "B625", "B623", "B901", "B1164", "B749", "B421", "B1183", CRC dataset size became 1196 (1203-7).

After consulting with Dr Snook by e-mail on 16th Jan 2019 about the records which read as “C189 - Malignant neoplasm of colon, unspecified” in the diagnosis: "B112" "B157" "B230" "B751" "B900" "B905" "B946" "B957" "B961"

Dr Snook has checked the records, and confirmed that:

1. "B112", "B230", "B751", "B946" were all correct in which histology for each entry is taken from metastasis. No CT or colonoscopy, or primary tumour were seen on CT.
2. "B900" histology was taken from peritoneal fluid, primary tumour was not seen on CT.
3. "B157" & "B957" were both incorrect: amended each one to malignant neoplasm of sigmoid colon in the diagnosis variable.
4. "B961" & "B905": No evidence of colonic primary, so advised to remove both records from the dataset.

After amending and removing the last two records, CRC dataset becomes 1194 (1196-2).

- After consulting with Dr Snook to confirm the TNM. Stages for patients IDs; “B79”, “B1046”, “B17”, and “B26”:

Dr Snook has checked the records, and confirmed that:

1. The stage for B79 should be changed from 3 to 2
2. The stage for B1046 should be changed to 2
3. The stage for B17 and B26: unstageable, as no scan or surgery

We excluded those with no tumour TNM. Stages from the final analysis. After taking out 17 records of TNM.Stage with missing values, CRC became 1177 (1194-17).

- After removing patients group C ; “other surveillance eg cancer follow-up, incidental finding on scan “because it did not fit in any other category, the dataset became 1143 (1177-34).
- After removing 5 records has no location, , the dataset became 1138 (1143-5).
- There were 17 entries with no Hb (these were left in).

Appendix III: Ethical approvals

Appendix IIIa: No Need for the NHS Approval

Go straight to content.

NHS
Health Research Authority

MRC Medical Research Council

Do I need NHS REC approval?

This decision tool suggests that you do not need NHS REC approval, however, you may still require another type of ethics committee review, e.g. Higher Education Institutions (HEIs) ethical approval.

Researchers in HEIs are advised to check whether, under their institution's policy and internal arrangements, ethical review is required by their HEI research ethics committee.

Exceptionally, the Research Ethics Service may accept an application for review of research at the request of the sponsor, chief investigator or host organisation, where it agrees that the proposal raises material ethical issues. Agreement should be sought from the responsible operational manager for the local REC centre prior to submission of the application.

Requests should be sent by email, including a summary of the research proposal (maximum one page) and explanation of why the project raises significant issues which cannot be managed routinely in accordance with established guidelines and good practice, and requires ethical consideration and advice from an NHS REC. Contact points for operational managers can be found on the HRA website.

Researchers requiring further advice (e.g. those not confident with the outcome of this tool) should contact their R&D office or sponsor in the first instance, or the HRA to discuss your study. If contacting the HRA for advice, do this by sending an outline of the project (maximum one page), summarising its purpose, methodology, type of participant and planned location as well as a copy of the previous results page and a summary of the aspects of the decision(s) that you need further advice on to the HRA Queries Line at HRA.Queries@nhs.net.

Follow this link to start again.

About this tool Feedback Contact Glossary

Appendix IIIb: BU Ethics Checklist Approval Decision

Bournemouth University

Research Ethics Checklist

About Your Checklist	
Reference Id	19925
Status	Approved
Date Approved	22/02/2018 09:37:53
Date Submitted	16/02/2018 13:02:19

Researcher Details	
Name	Orouba Almilaji
Faculty	Faculty of Health & Social Sciences
Status	Postgraduate Research (MRes, MPhil, PhD, DProf, EngD, EdD)
Course	Postgraduate Research - HSC
Have you received external funding to support this research project?	Yes
RED ID	10288
Funding body	: Poole Hospital NHS Trust
Please list any persons or institutions that you will be conducting joint research with, both internal to BU as well as external collaborators.	Poole Hospital NHS Trust

Project Details	
Title	The Importance of Identifying Iron Deficiency Anaemia in the Early Detection of Colorectal Cancer
End Date of Project	19/06/2020
Proposed Start Date of Data Collection	12/06/2004
Original Supervisor	Peter Thomas
Approver	Martin Hind

Summary - no more than 500 words (including detail on background methodology, sample, outcomes, etc.)

Background:A potential currently neglected approach to the earlier diagnosis of colorectal cancer (CRC) is through the detection of iron deficiency (IDA). IDA can result from long-term blood loss from a gastrointestinal lesion.It appears that IDA is a feature of a different subgroup of CRC than those who have a positive FaecalOccult Blood (FOB) test, and in particular most occur in the right colon (which carries a particularly poor outcome).Investigating the association between CRC and IDA, and identifying the subgroups of IDA patients who are at increased/lower risk of CRC might lead to speeding the investigation of CRC and thus reducing its potential prognostic implications, and helping patients avoid invasive procedures.Aims:1. Assess the prevalence of IDA in patients with CRC and estimate the duration of IDA prior to diagnosis.2. Assess the current frequency of blood count testing to detect IDA in at-risk age-groups.3. Assess whether cancers diagnosed through the detection of IDA are associated with an earlier tumour stage.4. Assess whether FOB testing can improve the clinical prediction of CRC in patients with IDA, and allow those at predicted low risk of CRC to safely avoid unnecessary

Appendix IIIc: Capacity and Capability at Poole Hospital

Confirmation of Capacity and Capability - The Importance of Identifying Iron Deficiency...

Reply | Reply All | Forward | ...

CS

Chessell, Sarah <Sarah.Chessell@poole.nhs.uk>
To Peter Thomas
Cc Orouba Almilaji; Crowley, Sacha; Burrows, Mary.T; Jonathon Snook (EX); Smith, Jayne; Younger, Margaret

Thu 03/05/2018 15:52

Dear Peter

RE: Confirmation of Capacity and Capability at Poole Hospital NHS Foundation Trust
Full Study Title: The Importance of Identifying Iron Deficiency Anaemia in the early detention of colorectal cancer

This email confirms that **Poole Hospital NHS Foundation Trust** has the capacity and capability to deliver the above referenced study.

If you wish to discuss further, please do not hesitate to contact me.

Kind regards,

Sarah Chessell

Sarah Chessell
Head of Research, Innovation, NICE and Clinical Audit
Cornelia House (1st Floor – Rm 32)
Poole Hospital NHS Trust
Direct line: 01202 448125 or email: sarah.chessell@poole.nhs.uk

Appendix IIId: Pilot Study (HRA approval)

NHS
Health Research Authority

Dr Jonathon Snook
Consultant Physician
Gastroenterology Department
Poole Hospital
Dorset
BH15 2JB
jonathon.snook@poole.nhs.uk

Email: hra.approval@nhs.net

07 October 2016

Dear Dr Snook

Letter of HRA Approval

Study title:	**The role of faecal occult blood testing in risk stratification for GI malignancy in subjects with iron deficiency anaemia**
IRAS project ID:	**201759**
Protocol number:	**P150915**
REC reference:	**16/LO/1464**
Sponsor	**Poole Hospital NHS Foundation Trust**

I am pleased to confirm that **HRA Approval** has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications noted in this letter.

Participation of NHS Organisations in England

The sponsor should now provide a copy of this letter to all participating NHS organisations in England.

Appendix B provides important information for sponsors and participating NHS organisations in England for arranging and confirming capacity and capability. **Please read *Appendix B* carefully**, in particular the following sections:

- *Participating NHS organisations in England* – this clarifies the types of participating organisations in the study and whether or not all organisations will be undertaking the same activities
- *Confirmation of capacity and capability* - this confirms whether or not each type of participating NHS organisation in England is expected to give formal confirmation of capacity and capability. Where formal confirmation is not expected, the section also provides details on the time limit given to participating organisations to opt out of the study, or request additional time, before their participation is assumed.
- *Allocation of responsibilities and rights are agreed and documented (4.1 of HRA assessment criteria)* - this provides detail on the form of agreement to be used in the study to confirm capacity and capability, where applicable.

Further information on funding, HR processes, and compliance with HRA criteria and standards is also provided.

It is critical that you involve both the research management function (e.g. R&D office) supporting each organisation and the local research team (where there is one) in setting up your study. Contact details and further information about working with the research management function for each organisation can be accessed from www.hra.nhs.uk/hra-approval.

Appendices

The HRA Approval letter contains the following appendices:

- A – List of documents reviewed during HRA assessment
- B – Summary of HRA assessment

After HRA Approval

The document *"After Ethical Review – guidance for sponsors and investigators"*, issued with your REC favourable opinion, gives detailed guidance on reporting expectations for studies, including:

- Registration of research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics, and is updated in the light of changes in reporting expectations or procedures.

In addition to the guidance in the above, please note the following:

- HRA Approval applies for the duration of your REC favourable opinion, unless otherwise notified in writing by the HRA.
- Substantial amendments should be submitted directly to the Research Ethics Committee, as detailed in the *After Ethical Review* document. Non-substantial amendments should be submitted for review by the HRA using the form provided on the HRA website, and emailed to hra.amendments@nhs.net.
- The HRA will categorise amendments (substantial and non-substantial) and issue confirmation of continued HRA Approval. Further details can be found on the HRA website.

Scope

HRA Approval provides an approval for research involving patients or staff in NHS organisations in England.

If your study involves NHS organisations in other countries in the UK, please contact the relevant national coordinating functions for support and advice. Further information can be found at http://www.hra.nhs.uk/resources/applying-for-reviews/nhs-hsc-rd-review/.

If there are participating non-NHS organisations, local agreement should be obtained in accordance with the procedures of the local participating non-NHS organisation.

IRAS project ID	201759

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please email the HRA at hra.approval@nhs.net. Additionally, one of our staff would be happy to call and discuss your experience of HRA Approval.

HRA Training

We are pleased to welcome researchers and research management staff at our training days – see details at http://www.hra.nhs.uk/hra-training/

Your IRAS project ID is **201759**. Please quote this on all correspondence.

Yours sincerely

Gemma Oakes
Assessor

Email: hra.approval@nhs.net

Copy to: *Ms Sarah Chessell, Poole Hospital NHS Foundation Trust [Sponsor Contact]*
sarah.chessell@poole.nhs.uk
Ms Margaret Younger, Research & Innovation Manager, Poole Hospital NHS Foundation Trust [Lead NHS R&D Contact]
margaret.younger@poole.nhs.uk

Appendix IIIe: MHRA Consultation with regard to the CE marking and Risk Class of the App

From: Hagan-Brown, Joe [mailto:Joe.Hagan-Brown@mhra.gov.uk]
Sent: 13 November 2018 12:21
To: Suzy Wignall
Cc: MHRA Customer Services
Subject: RE: GCEP-00132201 - FW: Query regarding device classification E/2018/1565

Dear Suzy,

Thank you for your enquiry to the MHRA dated 1st November 2018.

Please find answers to your questions below.

Please do not hesitate to contact me again if you require any additional information.

Kind regards,

Joe

Joe Hagan-Brown
Regulatory Affairs manager - Market Surveillance

MHRA, 10 South Colonnade, Canary Wharf, London E14 4PU
Email: joe.hagan-brown@mhra.gov.uk
Telephone: 0203 080 6786
Mobile: 07887452758
gov.uk/mhra

The app is intended to be used for diagnosis/prognosis of GI cancer. Apps which are specifically intended for the purpose of diagnosis/prognosis of disease are qualified as medical devices. The app is not intended to provide direct diagnosis (it is not intended that it itself provides a diagnosis independently), but rather it is intended the app will be an aid/decision support tool for the decision making that takes place between clinician and patient. MHRA's opinion, based on the information you have provided is that if the app is not providing direct diagnosis, and is not intended to provide decisive information for the diagnosis, then it will be qualified as an active device for diagnosis, which falls within class I under rule 12 of the annex IX classification rules of Directive 93/42/EEC (as amended). Please note that a higher classification (Class IIa likely) would apply under the new EU medical device regulation.

- if this *is* a medical device, due to the use of secondary anonymous data, how would the use of the device be approved? – would we be required to go through the IRAS application process for example;

The device would need to be CE marked as a medical device. Under the MDD, for class I medical devices CE marking guidance can be found here. An IRAS application would be required should you wish to conduct a clinical investigation of the non CE marked device. This may be necessary in the event that clinical data needs to be generated in order to demonstrate the safety and performance of the device, in accordance with annex X of the medical devices directive 93/42/EEC.
Please also be aware of the need to be prepared for the more stringent requirements introduced by the new EU regulations for medical devices. The commission has recently published guidance for manufacturers and I have provided a link below:
https://ec.europa.eu/growth/sectors/medical-devices/regulatory-framework_en

- the process required to attain CE marking for the app.

Please see the link above to the guidance for manufacturers of class I medical devices. For software based medical devices other MDD requirements which also apply include the need for the software to be validated (Essential requirement 12a of annex I of Directive 93/42/EEC).