

Article

The Development of a Clinical Decision-Support Web-Based Tool for Predicting the Risk of Gastrointestinal Cancer in Iron Deficiency Anaemia—The IDIOM App

Orouba Almilaji ^{1,2,*}, Vegard Engen ³, Jonathon Snook ² and Sharon Docherty ¹

¹ Department of Medical Science and Public Health, Bournemouth University, Bournemouth BH8 8GP, UK; sdocherty@bournemouth.ac.uk

² Gastroenterology Unit, University Hospitals Dorset NHS Foundation Trust, Poole BH15 2JB, UK; jonathon.snook@uhd.nhs.uk

³ Department of Computing and Informatics, Bournemouth University, Bournemouth BH12 5BB, UK; vengen@bournemouth.ac.uk

* Correspondence: oamilaji@bournemouth.ac.uk

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 and 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; IDIOM App



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. *Digital* **2022**, *2*, 104–119. <https://doi.org/10.3390/digital2010007>

Academic Editors: Andreas Triantafyllidis, Mariusz Duplaga and Konstantinos Votis

Received: 31 December 2021

Accepted: 16 March 2022

Published: 18 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

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, and 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 on 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 completed 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 no 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, such as 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.

Figure 1 illustrates the needed steps to put software medical device on market:

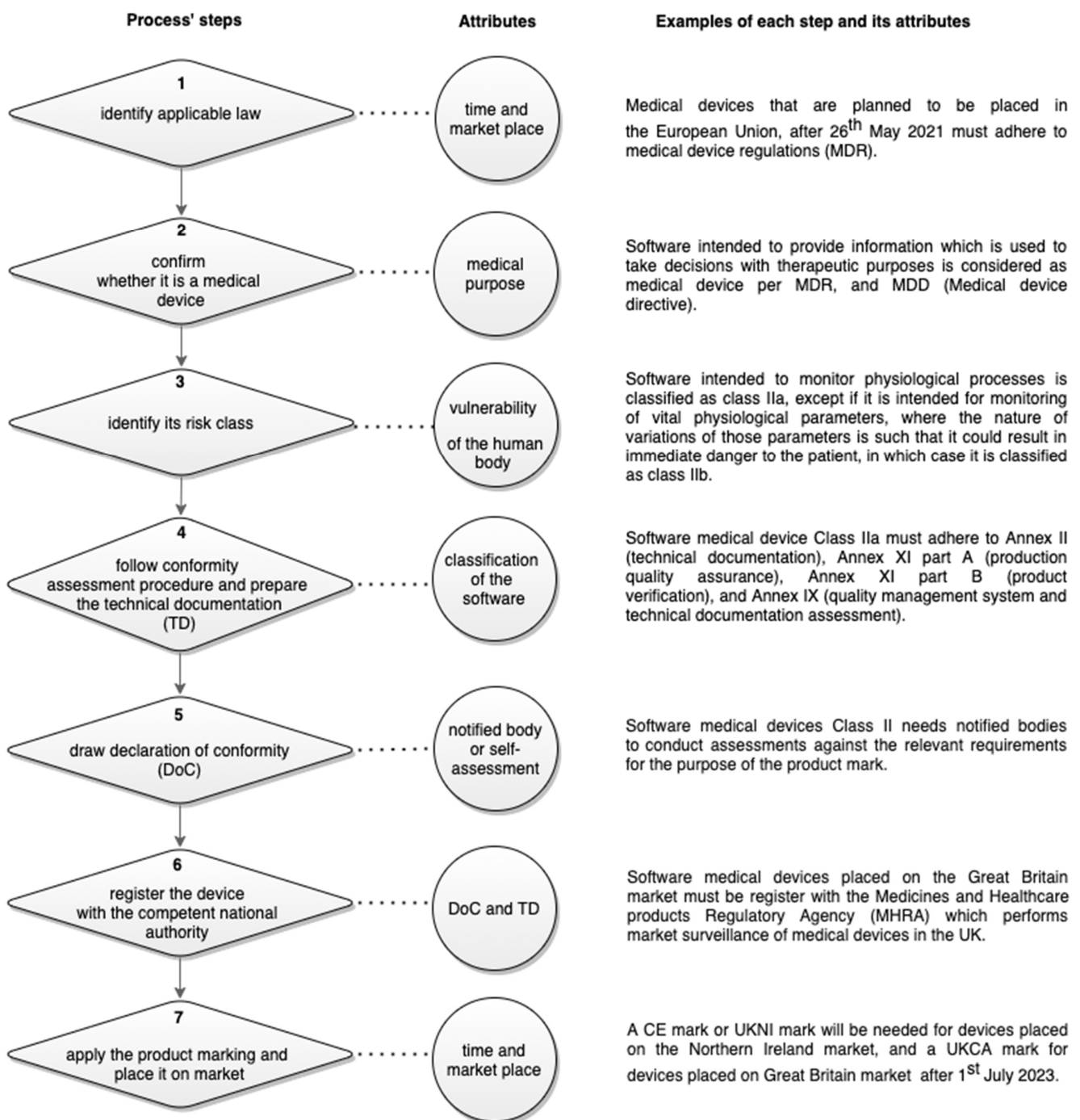


Figure 1. Steps of placing a software medical device on the market.

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. Additionally, 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 2, in which a very low Gi cancer risk is predicted.

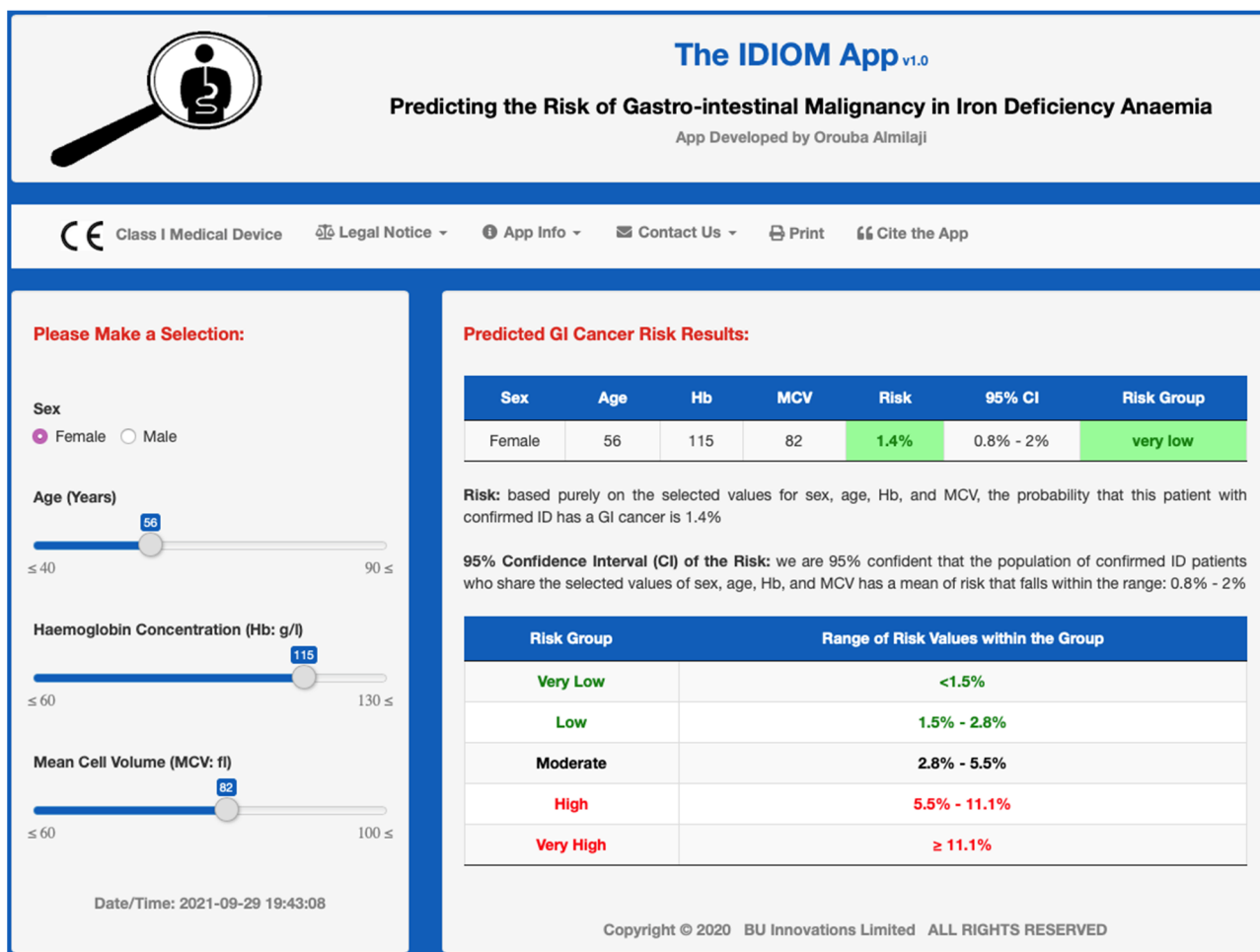


Figure 2. A screenshot of the IDIOM App.

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 and 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 current version of 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 3.

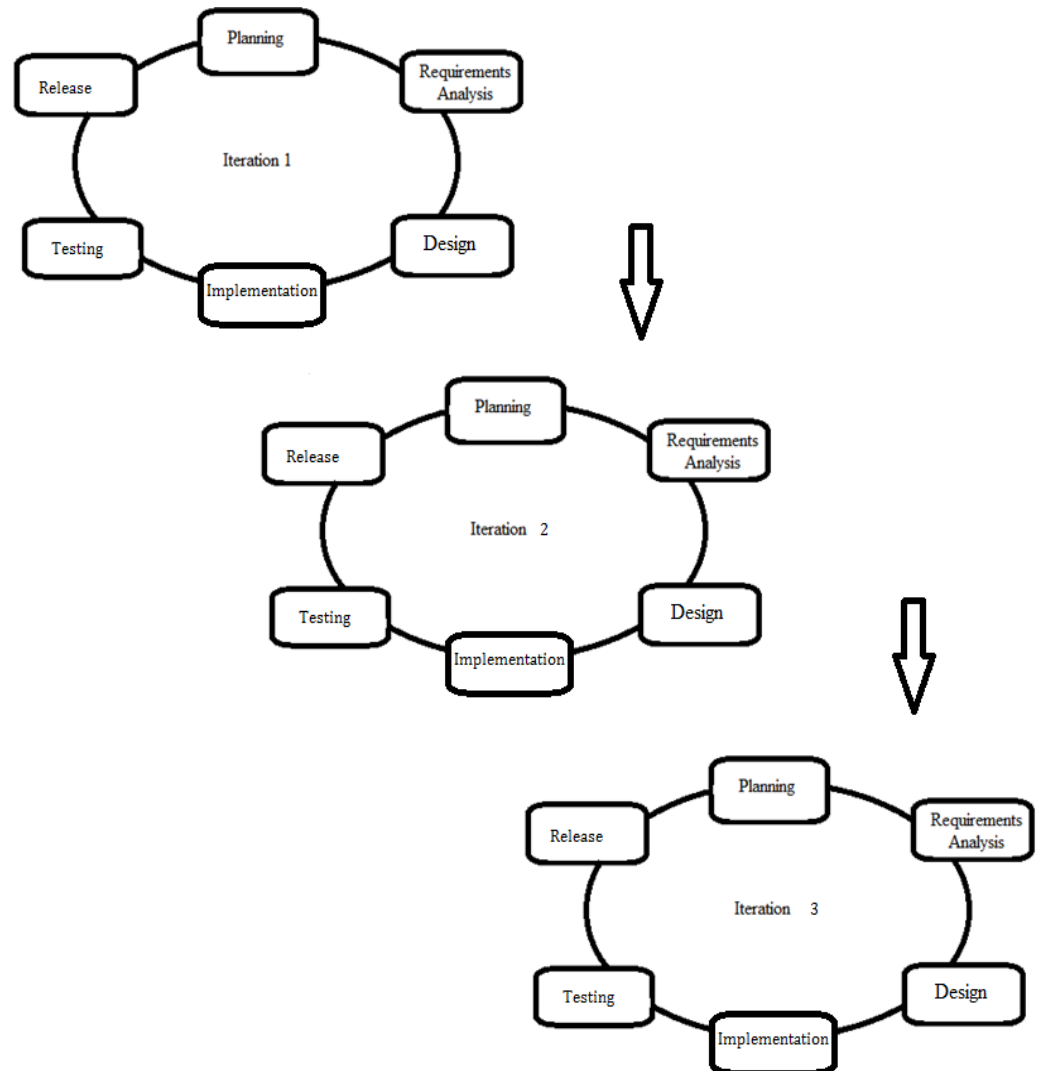


Figure 3. The agile model for developing the IDIOM App with three iterations, each iteration includes six tasks: planning, requirement analysis, design, implementation, testing, and release.

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 rationale of the app, and how to improve it [27–32].

- Other similar apps such as ‘predict prostate app’ [33] that has been developed at 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 and 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 the Moscow method [34] in which each requirement was categorized as (must have, should have, could have, or will not 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 as maintenance, compatibility, security, speed and performance, and overall control. The programming language which was used to 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 need 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 professionals who work usually in a 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. Additionally, the navigation between them was 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—Ph.D. candidate, O.A.—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. Additionally, 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. 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.

5.4.6. 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.

5.5. 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 [2,35–38]. 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 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 comprised 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]:

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

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

Table 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)

Density plots for the distribution of estimated risks per GI presence/absence using IDIOM model in each dataset is shown next in Figure 4.

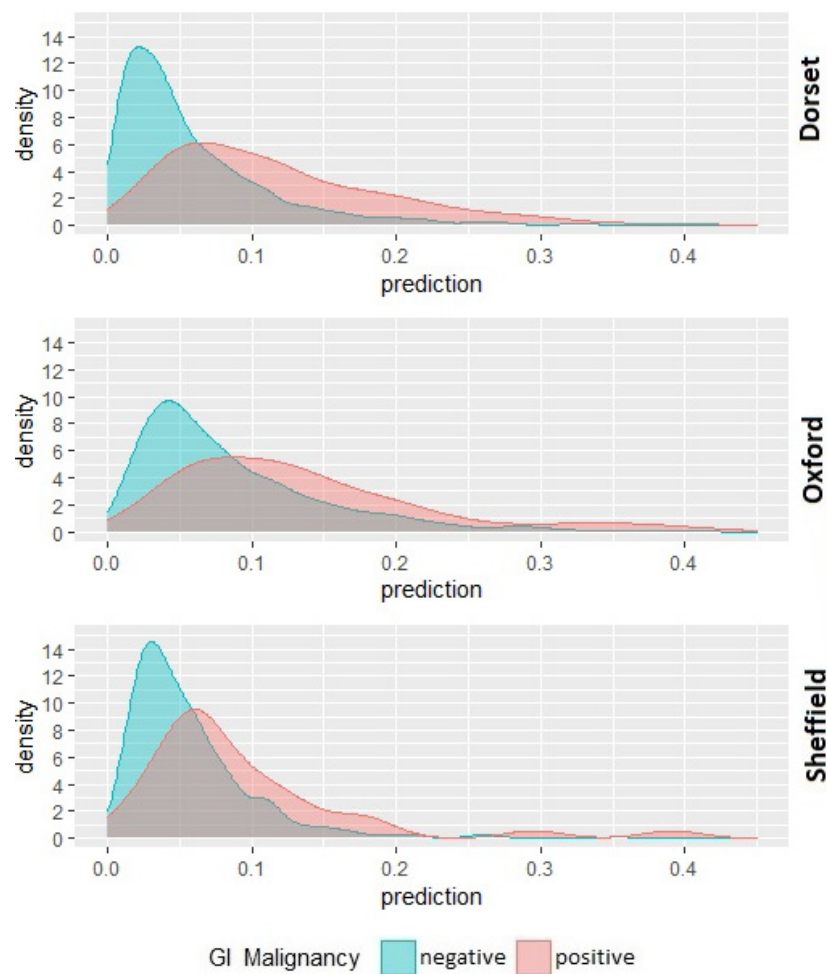


Figure 4. Density plots show the distribution of estimated risks per positive/negative GI in each dataset using IDIOM model.

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 χ^2 test, Akaike information criterion, and 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.

5.6. 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 undertaken at two points of time, where at the first time point, four participants assessed the interface usability. Additionally, 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 open feedback. Feedback generally 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.

5.7. 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;
- 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, and 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.

5.8. 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 fully 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 1 December 2020, with expected service life up to 30 June 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 the university to do things differently such as setting up surveillance system when the app went live, the IDIOM app was not distanced to succeed. The university 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 a 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 758. According to Google analytics, these users are from the UK (78%), Spain (4%), India (2%), Australia, Canada, Italy, USA, Malaysia, Portugal, Qatar, Thailand, Ireland, Saudi Arabia, Greece, South Korea, Mexico, New Zealand, Russia, Malta,

Colombia, Slovenia, Singapore, Poland, China, and other countries. Most of the access to the app comes through desktop (76%), then mobile (23%), and finally tablet machines (1%). Additionally, 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 stricter 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. Additionally, to discuss 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 performed according to the MDD regulations which were subsequently suppressed 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.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Acronyms, and 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

- 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. Risk of Cancer in Patients with Iron Deficiency Anemia: A Nationwide Population-Based Study. *PLoS ONE* **2015**, *10*, e0119647. [[CrossRef](#)] [[PubMed](#)]
- Silva, A.C.; Sheppard, Z.A.; Surgenor, S.L.; Williams, E.J.; Thomas, P.W.; Snook, J.A. Clinical risk factors for underlying gastrointestinal malignancy in iron deficiency anaemia: The IDIOM study. *Front. Gastroenterol.* **2014**, *5*, 237–242. [[CrossRef](#)] [[PubMed](#)]
- Stone, H.; Almilaji, O.; John, C.; Smith, C.; Surgenor, S.L.; Ayres, L.; Williams, E.J.; Snook, J. The dedicated iron deficiency anaemia clinic: A 15-year experience. *Front. Gastroenterol.* **2020**, *13*, 20–24. [[CrossRef](#)] [[PubMed](#)]
- Almilaji, O.; Smith, C.; Surgenor, S.; Clegg, A.; Williams, E.; Thomas, P.; Snook, J. Refinement and validation of the IDIOM score for predicting the risk of gastrointestinal cancer in iron deficiency anaemia. *BMJ Open Gastroenterol.* **2020**, *7*, e000403. [[CrossRef](#)] [[PubMed](#)]
- Almilaji, O.; Parry, S.D.; Docherty, S.; Snook, J. Colorectal Cancer and the Blood Loss Paradox. Available online: <https://fg.bmj.com/content/early/2021/10/20/flgastro-2021-101959> (accessed on 30 December 2021).
- Almilaji, O.; Parry, S.D.; Docherty, S.; Snook, J. Evidence for improved prognosis of colorectal cancer diagnosed following the detection of iron deficiency anaemia. *Sci. Rep.* **2021**, *11*, 13055. [[CrossRef](#)] [[PubMed](#)]
- National Institute for Health and Care Excellence (NICE). Suspected Cancer: Recognition and Referral. Available online: <https://www.nice.org.uk/guidance/ng12> (accessed on 30 December 2021).
- Almilaji, O.; Snook, J.; Thomas, P. *The IDIOM App*, version 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): Bournemouth, UK, 2020. Available online: <https://www.predict-gi-risk-in-ida.com> (accessed on 30 December 2021).
- Spiro, E. The Incidence of Time in the Conflict of Laws. *Int. Comp. Law Q.* **1960**, *9*, 357–382. [[CrossRef](#)]
- Medical Device Directive (MDD) EU 93/42/EEC of 14 June 1993 Concerning Medical Devices as Amended by Directive 2007/47/EC*; European Union: Brussels, Belgium, 2007.
- Medical Device Regulation (MDR) EU 2017/745*; European Union: Brussels, Belgium, 2017.
- IEC 62304; Medical Device Software—Software Life Cycle Processes*. The International Electrotechnical Commission: Geneva, Switzerland, 2006.
- IEC 62366; Medical Devices—Application of Usability Engineering to Medical Devices*. The International Electrotechnical Commission: Geneva, Switzerland, 2005.
- ISO 14971; Medical Devices—Application of Risk Management to Medical Devices*. The International Organization for Standardization: Geneva, Switzerland, 2019.
- ISO 9241-210; Ergonomics of Human-System Interaction: Human-Centered Design for Interactive Systems*. The International Organization for Standardization: Geneva, Switzerland, 2019.
- ISO 13485; Medical Devices Quality Management*. The International Organization for Standardization: Geneva, Switzerland, 2016.
- Implementation of Risk Management Principles and Activities within a Quality Management System*; The Global Harmonization Task Force (GHTF): Brussels, Belgium, 2005.
- Clinical Risk Management: Its Application in the Manufacture of Health IT Systems (DCB0129)*; Published under section 250 of the Health and Social Care Act 2012 and released on June 2018; NHS Digital Clinical Safety Team: Leeds, UK, 2018.
- Clinical Risk Management: Its Application in the Deployment and Use of Health IT Systems (DCB0160)*; Published under section 250 of the Health and Social Care Act 201 and released on June 2018; NHS Digital Clinical Safety Team: Leeds, UK, 2018.
- Classification of Medical Devices*; European Commission Health and Consumer Directorate B, Medical Devices Documents (MEDDEV 2. 4/1 Rev. 9); European Commission: Brussels, Belgium, 2010.
- 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: Brussels, Belgium, 2016.

22. *Medical Device Stand-Alone Software Including Apps*; The UK Medicines and Healthcare Products Regulatory Agency (MHRA): London, UK, 2014.
23. *Software as a Medical Device (SAMD): Key Definitions*; International Medical Device Regulators Forum (IMDRF): Brussels, Belgium, 2013.
24. *Software as a Medical Device (SAMD): Clinical Evaluation*; International Medical Device Regulators Forum (IMDRF): Brussels, Belgium, 2017.
25. *Principles of Labelling for Medical Devices and IVD Medical Device*; International Medical Device Regulators Forum (IMDRF): Brussels, Belgium, 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): London, UK, 2020.
27. Almilaji, O.; Engen, V.; Thomas, P.; Snook, J. O67 The development of a web-based application to predict the risk of GI cancer in IDA. *Gut* **2021**, *70*, A37–A38. [[CrossRef](#)]
28. Almilaji, O.; Thomas, P.; Snook, J. PWE-042 Predicting the risk of gastro-intestinal cancer in iron deficiency anaemia. *Gut* **2019**, *68*, A192. [[CrossRef](#)]
29. Almilaji, O.; Parry, S.; Thomas, P.; Snook, J. P288 Downstaging of right-sided colorectal cancer diagnosed through iron deficiency anaemia. *Gut* **2021**, *70*, A190. [[CrossRef](#)]
30. Almilaji, O. S1313 Modelling the Episodes of Care for IDA Patients in a Secondary Care Centre Using Continuous-Time Multistate Markov Chain. *Am. J. Gastroenterol.* **2021**, *116*, S605. [[CrossRef](#)]
31. 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 gastro-intestinal malignancy in iron deficiency anaemia. In Proceedings of the NCRI Cancer Conference, Online, 2–3 November 2020.
32. Almilaji, O. Modelling the episodes of care for iron deficiency anemia patients in a secondary-care center using continuous-time multistate Markov chain. *Saudi J. Gastroenterol.* **2022**, *28*, 115–121. [[CrossRef](#)] [[PubMed](#)]
33. Predict Prostate Homepage. Available online: <https://prostate.predict.nhs.uk/> (accessed on 31 August 2021).
34. Filipe, J.; Maciaszek, L.A. *Evaluation of Novel Approaches to Software Engineering*; Springer: Berlin/Heidelberg, Germany, 2013; ISBN 9783642540929/3642540929.
35. Capurso, G.; Baccini, F.; Osborn, J. Can patient characteristics predict the outcome of endoscopic evaluation of iron deficiency anemia: A multiple logistic regression analysis. *Gastrointest. Endosc.* **2004**, *59*, 766–771. [[CrossRef](#)]
36. Ho, C.-H.; Chau, W.-K.; Hsu, H.-C.; Gau, J.-P.; You, J.-Y.; Chen, C.-C. Predictive risk factors and prevalence of malignancy in patients with iron deficiency anemia in Taiwan. *Am. J. Hematol.* **2005**, *78*, 108–112. [[CrossRef](#)] [[PubMed](#)]
37. James, M.W.; Chen, C.-M.; Goddard, W.P.; Scott, B.B.; Goddard, A.F. Risk factors for gastrointestinal malignancy in patients with iron-deficiency anaemia. *Eur. J. Gastroenterol. Hepatol.* **2005**, *17*, 1197–1203. [[CrossRef](#)] [[PubMed](#)]
38. Wijayasekara, C.M.; Surgenor, S.L.; Hatcher, C.A.; Clegg, A.J.; Williams, E.J.; Snook, J.A. Clinical risk factors for underlying gastrointestinal malignancy in iron deficiency anaemia—prospective validation of the IDIOM score. *Cogent Med.* **2016**, *3*, 1213216. [[CrossRef](#)]
39. Altman, D.G.; Royston, P. The cost of dichotomising continuous variables. *BMJ* **2006**, *332*, 1080. [[CrossRef](#)] [[PubMed](#)]
40. Almilaji, O.; Webb, G.; Maynard, A.; Chapman, T.P.; Shine, B.; Ellis, A.J.; Hebden, J.; Docherty, S.; Williams, E.J.; Snook, J. Broad external validation of a multivariable risk prediction model for gastrointestinal malignancy in iron deficiency anaemia. *Diagn. Progn. Res.* **2021**, *5*, 23. [[CrossRef](#)] [[PubMed](#)]
41. Stone, D.; Jarrett, C.; Woodroffe, M.; Minocha, S. *User Interface Design and Evaluation*; Elsevier Science: Amsterdam, The Netherlands, 2005; ISBN 9780080520322/0080520324.
42. Snook, J.; Bhala, N.; Beales, I.L.P.; Cannings, D.; Kightley, C.; Logan, R.P.H.; Pritchard, M.; Sidhu, R.; Surgenor, S.; Thomas, W.; et al. British Society of Gastroenterology guidelines for the management of iron deficiency anaemia in adults. *Gut* **2021**, *70*, 2030–2051. [[CrossRef](#)] [[PubMed](#)]