CANDID
CANcer DIagnosis Decision rules

# PROTOCOL

**Version 4, dated: 3rd July 2015**

# CANcer DIagnosis Decision rules
# CANDID Study

**Funder:** NIHR National School for Primary Care Research (NSPCR)

Funding Reference: FR4/121
Amount: £2,085,433

Start Date: 01/03/2012
End Date: 01/01/2020

**Sponsor:** University of Southampton
Reference: 8310

**Chief Investigator**
Prof Paul Little
University of Southampton
Primary Care Group
Aldermoor Health Centre
Aldermoor Close
Southampton SO16 5ST
p.little@soton.ac.uk

**Co-applicants:**

(Co PI) Prof Tom Fahey, Royal College of Surgeons, Ireland

Dr Alastair Hay, University of Bristol

Dr Richard Stevens, University of Oxford

Dr Clare Bankhead, University of Oxford

Prof Richard Hobbs, University of Oxford

Professor Irwin Nazareth, University College London

Prof Danielle Van der Windt, Keele University

Dr Evangelos Kontopantelis, University of Manchester

Dr Gerry Leydon, University of Southampton

Dr Tom Sanders, Keele University

Dr William Hamilton, Peninsular Medical School, University of Exeter

Dr Colin McCowan, University of Dundee

Professor Frank Sullivan, University of Dundee

Professor Martin Dawes, University of British Columbia

Prof Brendan Delaney, Kings College, London

Prof Norbert Donner Banzhott, Philipps-Universitat Marburg

Prof Frank Buntinx, University of Leuven and Maastricht

Prof Paul Roderick, University of Southampton

Prof Tony Avery, University of Nottingham

Ms Samantha Hall, PPI representative

| **Glossary** | |
|---|---|
| | |
| CPRs | Clinical Prediction Rules |
| ESR | Erythrocyte sedimentation rate |
| CRP | C-reactive protein |
| CXR | Chest X-ray |
| PPV | Positive Prediction Value |
| haemoptysis | expectoration (coughing up) of blood or of blood-stained sputum from the bronchi, larynx, trachea, or lungs |
| dyspnoea | Subjective symptom of breathlessness |
| thrombocytosis | Presence of high platelet counts in the blood, and can be either primary or reactive |
| **PCRN** | **Primary Care Research Network** |
| | |
| | |
| | |
| | |
| | |

## Lay Summary

In primary care the key areas of concern for both doctor and patients are delay in diagnosing cancer, getting high risk patients referred first, and keeping investigation to a minimum. There have been few valid studies to assist decision-making in primary care, either to get a patient referred quickly or to assist in making sure an anxious patient is effectively reassured. This study seeks to work out which of the symptoms and examination findings are the most effective in predicting lung or colon cancer. To decide the best clinical information to collect in the study we will interview patients and also get consensus from a group of experts. Then we will recruit 20,000 patients who consult their GP - half with lung symptoms and the other half with low bowel symptoms. Clinical information will be collected using standardised internet based forms. Willing patients will complete lifestyle questionnaires and provide blood or saliva samples (including for genetic analysis). The National Cancer Registry will then be monitored to see which patients develop cancer, and statistical analysis will determine the most important clinical variables that predict cancer. The clinical prediction 'rules' or decision aids developed from these studies will then be tested with a further 2000 patients for each condition for validity.

**Brief overview of CANDID**

There have been very few studies to develop prospectively and then adequately validate clinical prediction rules (CPRs) for cancer in primary care, and yet concern about delaying a diagnosis of major pathology but avoiding over investigation remains a major concern for both patients and doctors at first presentation of symptoms in primary care.

This study involves eight departments of the NSPCR as recruitment hubs, and three departments have additional roles (coordination, statistical analysis, Delphi study). It is hoped that the study will also lead to the funding of several related studies – both new cohorts and studies of impact analysis.

**Background and rationale**

Observational studies based on routine data [1;2] have the great advantage of efficiently identifying possible 'signals' for cancer but given the limitations of possible differential recording of clinical data by GPs, such studies make it difficult to adequately quantify the importance of individual variables and their possible weighting – and so make it very difficult to develop valid CPRs. There is promising research for two of the commonest cancers seen in primary care (Lung and Colon) which suggest CPRs for these cancers should be possible but, again, there are significant limitations to these data.

**Lung cancer**

The 30,326 deaths from lung cancer in England and Wales in 2008 represent 22% of the total mortality burden from cancer (Office of National Statistics mortality data) - higher than either breast or colon cancer. Patients with lung cancer in the UK also present later and do worse than in other countries[3] - raising the issues of both prompt diagnosis and effective treatment.

Patients with lung cancer often recall having new symptoms frequently over the year before diagnosis – commonly cough, increasing shortness of breath or pain in the chest and also a cluster of systemic symptoms – e.g. fatigue/lethargy, weight loss, nausea vomiting, loss of appetite and altered taste change - which remain stable over time [4;5].

A systematic review has quantified the relative importance of individual symptoms[6] – in order of likely importance: haemoptysis, fatigue, cough, finger clubbing, weight loss, and dyspnoea - but none of these figures were derived from single primary care studies. A population based cohort study in routine data documented the relative importance of symptoms - loss of appetite (odds ratio 86), haemoptysis (32), dyspnoea (4.7), loss of weight (4.3) fatigue (3.2) thoracic pain (2.9), a second attendance with cough (2.7) - one physical sign (finger clubbing: 18), and two abnormal investigation results (thrombocytosis (9.3); abnormal spirometry (7.5).

However, quantifying the predictive values of symptoms and signs using routine data- bases is problematic: high positive predictive values are based on the symptoms GPs record (e.g. for haemoptysis the positive predictive value (PPV) from routine data bases is likely to be around 7.5% (6.6% to 8.5%)[7], but the PPV - particularly among younger patients - is very likely to provide a significant over-estimate not only because routine data bases rely on the GP's discretion on the choice of symptoms to code, but also because GPs are more likely to document symptoms that might suggest cancer if they intend to refer than if they don't (for example a small amount of blood mixed with sputum in a young patient with a presumed chest infection is less likely to be coded). NICE guidelines suggest that any haemoptysis, or any of the above symptoms lasting longer than three weeks should be investigated with a chest x-ray[8] but we know that for the commonest acute RTI presenting in primary care the median duration of symptoms is 3 weeks[9;10] so this guidance arguably may be setting too low a threshold for investigation. There is also evidence from secondary care settings that a normal x-ray may not be helpful in excluding cancer [11].

If clinicians in primary care acted on the NICE guidance for x-rays this could dramatically increase the number of chest x-rays performed for the primary care population; whilst this may be appropriate, its may also increase the dangers of iatrogenesis, and may not be a cost-effective approach to diagnosis. A clinical prediction rule based on prospective clinical data collection and assessing the place of simple investigations in primary care (full blood count, and chest x-ray) is the most robust way to better inform thresholds for such investigations.

**Colorectal cancer**

Colorectal Cancer is the third most commonly diagnosed cancer in the UK with almost 40,000 new cases per year documented in 2007 (Office of National Statistics), similar to lung cancer, and second only to lung cancer in mortality. Patients with rectal bleeding commonly present in primary care[12] , but only 2-11% are likely to have serious disease[13;14]. NICE recommends urgent referral for those with rectal bleeding if aged 40 years or older and persistent looser stools and/or increased stool frequency, or alternatively if aged 60 and older with isolated rectal bleeding or with persistent changed bowel habit without anal symptoms (NICE, 2005).

As with lung cancer, referring patients at low risk of colorectal cancer may lead to patient anxiety and iatrogenesis from further diagnostic investigations and longer waiting time for high-risk patients. A systematic review of 8 diagnostic studies with 2323 patients with rectal bleeding[15] identified age >60, weight loss, a change in bowel habit and anaemia all increase the probability of cancer. However, most of the studies were underpowered, the selection of potential predictors and also reference standards was variable, and the nature of diagnostic meta-analysis – the need to use univariate data – makes it unclear what variables are likely to be important in an adequately powered multivariate analysis. Furthermore such data needs to be generated in primary care.

A wider ranging systematic review of diagnostic studies from our group[16] suggests family history, weight loss, and iron deficiency anaemia are likely to be important but insufficiently studied in primary care, and a further systematic review from our group examined all

symptoms of colorectal cancer using only primary care data and has just been completed (Astin et al, BJGP in press) - and will help inform the selection of variables for our proposed cohort.

Two colorectal scores have been developed, but have not entered routine clinical practice: the SELVA score – derived in a surgical clinic setting [17] was only moderately useful when tested in a second referred population.[18] A scoring system based on routinely collected data in primary care (CAPER) has been developed , feasibility tested,[1;2] and in a second dataset has performed more favourably than the NICE algorithm (Marshall et al, Gut, in press). However, the CAPER score was developed based on routine data, and so the key issues regarding the validity of weighting of variables and the possible bias of missing variables from routine data sets applies. Therefore, here again prospective development and validation of a CPR in primary care is needed.

We are aware of a primary care cohort to study abdominal symptoms that has been set up by Norwegian investigators (lead by Knut Holtedahl) which relates to a much wider range of abdominal cancers. In contrast we propose concentrating on colorectal cancer to focus data collection, optimise diagnostic performance, maximise feasibility, and reduce the barriers to recruitment. Nevertheless , there are potentially important synergies from the Norwegian led study and this study: first, in more securely defining the evidence (since one study is rarely sufficient, and patient populations in different health care systems are likely to present differently leading to variations in diagnostic performance); secondly in cooperating to enable the mutual use of both data sets. Thus we will liaise with the Norwegian team prior to commencing data collection in order to ensure as much overlap of the relevant parts of the data proformas as is appropriate - with a view to being able to share data sets for testing the scores that are developed in each data set. In addition to the clinical presentation we will assess the increase in diagnostic performance when adding information from additional measures (e.g. genetic, inflammatory and lifestyle information including smoking and alcohol status) to prediction models based on symptoms and signs only.

**Objectives:**

1) We will develop web based clinical proformas for cohort studies based on prior systematic reviews, patient interviews, and a Delphi exercise to confirm candidate variables
2) Prospective diagnostic cohorts will be used to develop and validate Clinical Prediction Rules for lung and colon cancer
3) The incremental utility of incorporating additional measures (e.g. genetic, inflammatory and lifestyle information including smoking and alcohol status) in the prediction models will be explored.

**Methods**

We will generate two prospective cohorts of patients presenting with lung and colonic symptoms in order to develop clinical prediction rules for both lung and colon cancer. Based on the literature to date, interviews with patients, and a Delphi exercise we will develop and implement simple web based clinical proformas and then follow up patients in National Cancer Registries and GP records to ascertain cancer cases.

Patients will enter the study period within 4 weeks of the date they present to their general practitioner with lung or lower bowel symptoms consistent with those identified for inclusion in the appropriate cohort. Patients will also be invited to provide additional measures (e.g. genetic, smoking, dietary, and alcohol history) - but these will be optional to ensure no effect on recruitment and the most generalisable sample possible.

**Lung cohort**

Inclusion criteria

- Any adult patient > 35 years presenting with symptoms lasting for 3 weeks that could be associated with lung cancer – either focal chest symptoms (haemoptysis, dyspnoea, thoracic pain, cough) or systemic symptoms lasting for 3 weeks with no other localising symptoms (e.g. loss of appetite, loss of weight, fatigue)

Exclusion criteria

- Exclusion criteria: Known lung cancer, pregnancy, or urgent admission to hospital (e.g. massive haemoptysis), other terminal illness. Inability to provide a good history (severe depression, psychosis, dementia, acute alcohol intoxication, learning impairment)

**Colorectal cohort**

Inclusion criteria

- Inclusion criteria: Adults > 35 years (since colon cancer is very rarely diagnosed in the younger age group) presenting with lower gastrointestinal symptoms that could be associated with colorectal cancer. This includes any of the following symptoms: rectal bleeding, bowel symptoms (change in bowel habit, tenesmus, urgency, incomplete emptying, nocturnal symptoms) systemic symptoms (weight loss, anorexia, fatigue) lower abdominal pain.

Patients who have been included in the national screening programme may also be referred to the study.

Exclusion criteria

- Exclusion criteria: Known colorectal cancer, pregnancy, or urgent admission to hospital (massive bleeding or acute abdomen), other terminal illness. Inability to provide a good history (severe depression, psychosis, dementia, acute alcohol intoxication, learning impairment)

**Patient interviews and Delphi exercis**e

The list of variables to be included for each cohort will be agreed following a series of patient interviews - with patients diagnosed with lung, or colon cancer within the previous year - and then a Delphi exercise. (See Appendix 1 for details of the qualitative study).

**Recruitment of sites and participants**

Based on the literature, it is assumed that GPs will see one of the index conditions approximately 1-2 times per month and that the majority of patients will agree to participate based on the experiences of previous studies.

It is currently estimated that each of the 8 academic sites will need to identify 60-70 GP practices providing approximately 200 GPs.

**Academic Sites**

University of Southampton
University of Bristol
University of Manchester
University of Oxford
University of London
University of Nottingham
University of Birmingham
University of Keele

**Consent**

Potential participants will have their baseline data entered into the on line pro-forma. They will be given a participant information sheet and a consent form, and will be given time to take the information away and consider if they wish to take part. If they agree, they will be asked to return the consent form to the study team who can then include the baseline data already entered on to the website. If patients are happy to sign the consent form at the time of consultation or soon after the consultation they can do so and return the form to the recruiting Healthcare Professional who will forward it to the study team.

**Data collection for the cohorts**

Patients presenting with symptoms that could be indicative of a future diagnosis of cancer in primary care will have a structured clinical examination using standardised web based proformas. Patients will also be invited to provide additional measures (e.g. genetic, inflammatory and lifestyle information including smoking and alcohol status) - but these will be optional to ensure no effect on recruitment and the most generalisable sample possible.

Provisionally therefore we propose the following variables:

- **Lung cancer**

The variables included in the clinical prediction rule will be classified and considered in 4 separate groups

1) socio-demographic – age, sex, social class, Townsend deprivation score, ethnicity, family history of lung cancer, smoking. Socio-economic measures at baseline alongside other predictors (e.g. other major co morbidities, and if feasible health literacy) are likely to be relevant to initial presentation and referral (particularly delay), and may modify the impact of predictor variables.
2) symptoms – either focal chest symptoms (haemoptysis, dyspnoea, thoracic pain, cough) or systemic symptoms (loss of appetite, loss of weight, fatigue)
3) examination findings – focal chest signs
4) Further investigations for willing patients – full blood count, CRP, (see below).

- **Colorectal cancer**

The variables included in the clinical prediction rule will be classified and considered in 4 separate groups based on a local modification of the UK Department of Health guidelines:[21]

1) socio-demographic – age, sex, social class, ethnicity, family history of colorectal cancer, past history of type 2 diabetes, inflammatory bowel disease or benign polyps.
2) symptoms – rectal bleeding (type, duration, mixed with stool), bowel symptoms (change in bowel habit, tenesmus, urgency, incomplete emptying, nocturnal symptoms) systemic symptoms (weight loss, anorexia, fatigue), abdominal/anal pain.
3) examination findings – abdominal mass, rectal examination (unless declined by patient or clearly painful perianal condition e.g. anal fissure) lymphadenopathy
4) Further investigations for willing patients – full blood count, CRP, ferritin (see below).

GPs will be asked to document their reason for referral and also to estimate of the risk of cancer. For participants recruited in secondary care settings referral information will be used when completing the CRF.

Other measures: for both cohorts blood or saliva will be taken in some patients to provide samples to store for genetic analysis, also CRP, FBC, ferritin; willing patients will also be asked to complete validated web based questionnaires[19;20] (fruit and veg, exercise, family history).

Participants without access to the Internet will be given the opportunity to request a paper questionnaire. These variables are optional to minimise recruitment bias.

**Collection and Storage of samples**

**Saliva**

If the participant does not wish to provide a blood sample but is willing to provide a saliva sample then they may do this during the initial consultation with the GP or make another appointment with the nurse.

Once the sample has been taken, it will be sent via Royal Mail to the University of Southampton. It will be stored in a freezer in the Faculty of Medicine Human Tissue Bank until funding can be secured to allow DNA analysis of this.

**Blood**

Participants consenting to provide a blood sample will need to make an appointment with the practice nurse or appropriate health care professional. The samples (two clotted and two EDTA) will be sent via Royal Mail to the University of Southampton. A full blood count will be undertaken and the remainder will be stored in freezers in the Faculty of Medicine Human Tissue Bank until funding can be secured to analyse them.

The samples will be stored in HTA licensed premises and only released to appropriate research individuals in line with the Tissue Bank's Standard Operating Procedures.

**Outcome measures**

Diagnosis of colorectal or lung cancer though linkage to National Cancer Registries or GP records due to colorectal or lung cancer within two years to five years of presentation. The implications of using a 1 year cut-off will also be explored. Information on the stage and grade of cancer will be collected for descriptive purposes and for exploratory subgroup analyses. It could also provide a platform for a possible subsequent study to examine factors which predict survival following diagnosis of cancer.

**Proposed sample size**

**For the development cohort:** The study will recruit a minimum of 7200 patients but preferably 10,000 patients for each cohort. (See Table 1)

**Validation cohorts.** For the validation cohorts we aim to include a minimum of 2000 patients for each cohort. This will supply at least 100 cases for each cohort - as recommended from simulation studies[21].

Table 1. Numbers of controls/cases needed for development cohort using different assumptions

| | | Control n/ Case n | | | |
|---|---|---|---|---|---|
| | Control prevalence of risk factor | Alpha 0.05 Beta 0.2 | Alpha 0.01 Beta 0.2 | Alpha .05 Beta 0.1 | Alpha 0.01 Beta 0.1 |
| | | | | | |
| | 5% | 4748/238 | *6852/343* | 6699/335 | ***9161/459*** |
| | 10% | 2653/133 | 3848/193 | 3712/186 | 5105/256 |
| | 50% | 1436/72 | 2149/108 | 1902/96 | 2712/136 |
| | 80% | 3741/188 | 4426/222 | 3741/188 | 5441/273 |

**Statistical analysis.**

Logistic regression will be used to develop multivariate models to predict diagnosis of each cancer. Multiple imputation will be used if appropriate to replace missing values and fractional polynomials will be used to model non linear risks relations with continuous variables. We will examine for interactions between age, measures of socio-economic status, and each explanatory variable. The analysis will generate variables significant in multivariate analysis which will be included in the CPRs. We anticipate that only variables that have odds ratios of 2 or 0.5 in multivariate analysis are likely to be useful in a CPR. We propose exploring the role of repeated presentation of symptoms since it may be that a symptom has greater predictive power if presented for a second time.

We will develop a rule a) preferably based on the simple count of predictive variables (like the widely used Centor criteria for sore throat[22]) - which will be the simplest to use in clinical practice, and also b) we will also explore the possibility of using a weighted rule (where each variable is weighted according to the rounded logistic coefficients). The performance of the CPR in both the development cohort and in the validation set will be explored by assessing the sensitivity, specificity and predictive values at each cut point in the CPR. We will also use serial ROC curves - adding sequentially variables from the CPR, starting with the most predictive variable – in order to explore the likely utility of increasing the number of variables in the CPR, since the simplest CPR is the one that is most likely to be used. We will also assess discrimination using the R2 and D statistic, and assess classification/reclassification using the methods of Pencina et al[23] . We will assess calibration by comparing observed and predicted risks in each tenth of predicted risk. Similar methods will be used to investigate the added performance of including genetic, inflammatory or lifestyle information in the prediction rules.

The performance of all the CPRs developed to date will also be compared in the available data sets. This will obviously have limitations given the differing variables in each data set, but will provide some cross validation, and will allow a preliminary exploration of the likely influence of spectrum bias in the different settings.

**Minimising selection bias**:

1) In one of the Networks we propose exploring the possibility of developing an automated system within the clinical computer systems to prompt GPs when index symptoms present. We anticipate this will reduce selection bias, and improve recruitment, and if proven acceptable, feasible and timely will use this approach in other Networks.

2) Willing practices will be asked to perform an audit every 2 weeks of those presenting in the practice and invite eligible patients who have not been recruited into the study; the characteristics of patients recruited by this mechanism and by opportunistic recruitment will be compared.

**Recruitment rate.**

Based on the above literature we assume that GPs will see one of the index conditions approximately 1-2 times per month and that the majority of patients will agree to participate based on the experience of the DESCARTE Study. We anticipate we will need 60-70 practices, with 200 GPs, per centre.

Reference List


(1) Hamilton W. Derivation of a score for identifying colorectal cancer in primary care. *Gut* 2007;56(Suppl 2):A49. Gut 2007; 56 (s2):A49.

(2) Khan N, NCRI Colorectal Clinical Studies Group. Implementation of a diagnostic tool for symptomatic colorectal cancer in primary care: a feasibility study. Primary Care Research and Development 2009; 10:54-64.

(3) Hubbard R, Baldwin D. Diagnosing lung cancer earlier in the UK. Thorax 2010; 65:756-758.

(4) Corner J, Hopkinson J, Fitzsimmons D, Barclay S, Muers M. **Is late diagnosis of lung cancer inevitable? Interview study of patients' recollections of symptoms before diagnosis**. Thorax 2005; *60 doi:10.1136/thx.2004.029264*:314-319.

(5) Gift A, Stommel M, Jablonski A, Given W. A Cluster of Symptoms Over Time in Patients With Lung Cancer. Nurs Res 2003; 52:393-400.

(6) Hamilton W, Sharp D. Diagnosis of lung cancer in primary care: a structured review. Family Practice 2004; 21:605-611.

(7) Jones R, Latinovic R, Charlton J, Gulliford M. Alarm symptoms in early diagnosis of cancer in primary care: cohort study using General Practice Research Database. BMJ 2007; 334:*1040 doi: 101136/bmj.39171.637106.AE* .

(8) NICE (National Institute of Clinical Excellence). The diagnosis and management of lung cancer. NICE 2011; CG121 (replacing CG24).

(9) Little P, Rumsby K, Kelly J, Watson L, Moore M, Warner G et al. Information leaflet and antibiotic prescribing strategies for acute lower respiratory tract infection: a randomised controlled trial. JAMA 2005; 293:3029-3035.

(10) Moore M, Little P, Rumsby K, Kelly J, Watson L, Warner G et al. Predicting the duration of symptoms in lower respiratory tract infection. Br J Gen Pract 2008; 58:88-92.

(11) Thirumaran M, Sundar R, Sutcliffe I, Currie D. **Is investigation of patients with haemoptysis and normal chest radiograph justified?** Thorax 2009; 64:854-856.

(12) Chaplin A, Curless R, Thomson R, Barton R. Prevalence of lower gastrointestinal symptoms and associated consultation behaviour in a British elderly population determined by face-to-face interview. BJGP 2000; 50:798-802.

(13) Du Toit J, Hamilton W, Barraclough K. Risk in primary care of colorectal cancer from new onset rectal bleeding: 10 year prospective study. BMJ 2006; 333:69-70.

(14) Douek M, Wickramsinghe M, Clifton M. Does isolated rectal bleeding suggest colorectal cancer? Lancet 1999; 354:393.

(15) Olde Bekkink M, McCowan C, Falk G, Teljeur C, Van de Laar F, Fahey T. Diagnostic accuracy systematic review of rectal bleeding in combination with other symptoms, signs and tests in relation
to colorectal cancer. British Journal of Cancer 2010; 102:48-58.

(16) Jellema P, Van der Windt D, Bruinvels D, Mallen C, vna Weyenberg S, Mulder C et al. Value of symptoms and additional diagnostic tests for colorectal cancer in primary care: systematic review and meta-analysis. B M J 2010; *340:c1269* .

(17) Selvachandran S, Hodder R, Ballal M, Jones P, Cade D. Prediction of colorectal cancer by a patient consultation questionnaire and scoring system: a prospective study. *Lancet* 2002;360:278-83. Lancet 2011; 360:278-283.

(18) Ballal M, Selvachandran S, Maw A. Use of a patient consultation questionnaire and weighted numerical scoring system for the prediction of colorectal cancer and other colorectal pathology in symptomatic patients: A prospective cohort validation study of a Welsh population. Colrectal Disease 2010; 12:407-414.

(19) Little P, Barnett J, Margetts B, Kinmonth AL, Gabbay J, Thompson R et al. Validity of dietary assessment in general practice. J Epidemiol Community Health 1999; 53:165-172.

(20) Jacobs D, Ainsworth B, Hartman T, Leon A. A simultaneous evaluation of 10 commonly used physical activity questionnaires. Med Sci Sports Exercise 1993; 25:81-91.

(21) Vergouwe Y, Steyerberg E, Eijkemans M, Habbema D. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol 2005; 58:475-483.

(22) Centor RM, Witherspoon JM, Dalton HP. The diagnosis of strep throat in the emergency room. Med Decis Making 1981; 1:239-246.

(23) Pencina M, Agostino R, Agostino RJ, Ramachandran S. Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. Statistics in Medicine 2007;DOI: 10.1002/sim.2929.

Appendix 1. Qualitative study

**Lung and colorectal cancer patients' perceptions of 'non-classical' symptoms prior to presentation in primary care: a qualitative interview study**

**Aims:** To investigate patients' experiences and interpretation of 'non-classical' symptoms prior to a first consultation in primary care

**Research question:** How do lung and colorectal cancer patients interpret 'non-classical' symptoms prior to seeking healthcare advice

**Purpose and context:** We propose to conduct a series of qualitative semi-structured interviews with patients diagnosed with lung or colorectal cancer within the previous 12 months. The interviews will be conducted during the course of 9 months, alongside the development of the Cancer Diagnosis Decision Rules study (CANDID). The CANDID study will generate two prospective cohorts of patients presenting with lung and lower bowel symptoms in order to develop clinical prediction rules for both lung and colorectal cancer. Partly based on the qualitative interviews with patients, the CANDID study will develop and implement simple web based clinical proformas and then follow up patients in National Cancer Registries to ascertain cancer cases. The proformas will also be based on the results of a Delphi exercise, designed to build consensus between clinicians and researchers on potential diagnostic indicators (symptoms and signs) and tests considered crucial in the identification of patients presenting in primary care with symptoms indicating an increased risk of lung or colorectal cancer. Consequently, the findings from the interviews with patients will inform and feed into the Delphi consensus study and subsequently into the design of the larger CANDID cohort study.

The proposed qualitative interview study was funded as part of the CANDID cohort study, to inform the development of the prediction rules, and to identify the presence of 'non-classical' symptoms prior to presentation and diagnosis, as these have not been reported in the literature. The central question in the patient interviews will be: 'what happened or what did you feel or observe during the last weeks or months before diagnosis which was different from normal?' The main aim of this interview phase, to generate the key information relatively rapidly, is not only to provide rich contextual understanding of the path to presentation, but particularly to unearth non classical symptoms and explore these in detail and in relation to each patient's circumstances and social context. In addition, the study will develop understanding of how best to communicate 'no need for referral' decisions and reassurance to those patients at a lower risk of lung and colorectal cancer. Members of our group used a similar approach in a diagnostic study for serious diseases in children, and the approach proved invaluable in providing important indicators that would not have been included had the interviews not been performed. Exploration of people's symptoms prior to a cancer diagnosis will provide insights into how and why they responded to them in the way they did. Drawing on the concept of 'containment' (a conceptual framework for understanding symptom interpretation as a process that is grounded in day-to-day situations), we will

identify the context of patients' perceptions of 'non-classical' symptoms, going beyond simple 'recognition' of those symptoms. Patient interviews will address the role of relatives or other significant members of respondents' social networks and their contribution to shaping the patient's understanding of their symptoms and 'triggers' to consultation. Interviews will be guided using an interview schedule so that the main questions are addressed with all respondents, but allowing for individual flexibility in responses to questions. The interviews will focus on 'non-classical' symptoms that are reported by respondents or which we have identified in the literature that do not appear in the NHS guideline. The interviews will then explore in greater depth the specific symptoms, their meaning for patients, the role of their social context, and their relation to the patients' presentation pathway. The data will be analysed inductively (hypothesis generation) to identify common patterns and meanings in the respondents' accounts, and 'non-classical' symptoms will be identified (contextually) and used to inform the Delphi exercise.

**Study sample and recruitment:** A sample of 10 lung cancer (LC) and 20 colorectal cancer (CC) patients, diagnosed in the previous 12 months, will be selected purposively from a respiratory clinic and a colorectal cancer clinic respectively at the Leighton Hospital, North Staffordshire. We have chosen to sample 10 rather than 20 patients with lung cancer as colleagues from the University of Southampton have recently conducted a similar interview study with a larger number (28) of patients with suspected (8) or histologically confirmed (20), respectable lung cancer. The data set that we intend to build on was collected by one of the CANDID study co-investigators (LB) from the University of Southampton, and we have been given full access to these data to permit analysis. We will compare this dataset with our own findings, and extend the analysis to include additional topics not covered previously, such as how best to communicate 'no need for referral' and 'level of lung cancer risk'. We opted to recruit patients from secondary care rather than through general practice, in order to obtain the desired number of patients diagnosed in the previous 12 months. A large number of general practices would otherwise have to be approached in order to accrue similar patient numbers. Instead, recruitment from secondary care will offer a much larger pool of patients from which to purposively sample (see selection criteria below). We will approach approximately 50 patients (30 CC and 20 LC) with a confirmed clinical diagnosis. Based on an estimated participation rate of 60% we anticipate that we will achieve our sample of 30 patients (20 CC and 10 LC). Patients will be selected from a range of socio-economic backgrounds to provide a diverse mix of respondents. Patients will be selected according to the following criteria: urban/rural; 'younger' and 'older' patients (e.g. >50 and <50 years old); equal proportions of males and females; and different disease stage (equal proportion of early and more advanced stage disease).

The researcher will attend outpatient respiratory and colorectal cancer clinics (surgical) to approach all potential patients in person. Patients who are attending the clinic for a first or follow-up appointment, and who have been diagnosed in the previous 12 months, will be approached by a nurse and invited to talk directly with the researcher in a private room at the clinic, following their clinic appointment. Prior experience by researchers at the University of Southampton revealed that this method of recruitment led to a very high participation rate.

The researcher will explain the purpose of the study, and the interview process, and verbally invite them to take part. Patients who indicate that they do not want to participate in the interview at this stage will not be approached again. All patients verbally consenting to further contact about the interview will then be given a letter of invitation, and an information sheet, and asked to return a reply slip indicating if they want to be contacted again (using a pre-paid envelope). All patients consenting to an interview will then be contacted by telephone to discuss any questions and arrange a suitable time for the interview. The word 'cancer' will not be used in the study information to avoid causing potential distress, and to avoid labelling patients as having a cancer diagnosis in case they prefer not be identified in this way, or their diagnosis has not been histologically confirmed.

All patients who agree to participate will be offered the option to have a relative present during the interview (for support). The purpose will be explained to the patient again immediately before the interview, and written consent obtained prior to, and immediately after, the interview, to give respondents the opportunity to review their consent in light of the discussion. Patients will be told that they can withdraw from the study at any time prior to the appointment and they can stop the interview at any point. The emotive nature of the subject matter means that safeguards need to be introduced in case patients become upset, and may require further support from their oncologist and/or GP. Consequently, all GPs whose patients are interviewed will be contacted by letter and informed about the interview study. The researcher will seek agreement from the patient to contact their GP, prior to the interview. Patients will also be informed that if they become upset as a result of the interview or feel the need to discuss any aspect of their health or healthcare, they should approach their GP.

**Data collection:** Semi-structured qualitative interviews will identify reported 'non-classical' symptoms which may indicate a diagnosis of lung and colorectal cancer. The aim of the qualitative study is to identify key 'non-classical' symptoms, which will be used in part to inform the development of the Cancer Diagnosis Decision Rules [see above]. The reasoning is that primary care physicians are currently either unaware of non-classical symptoms, or may be uncertain how to incorporate them into decisions about referral for secondary assessment.

The interviews will generate in-depth data on patients' perceptions and experiences of non-classical symptoms prior to presentation to a GP. Respondents will also be interviewed about contextual characteristics such as the presence of a family history of cancer; symptoms on presentation to primary care, such as rectal bleeding or pain; and their experiences of the clinical examination and the disclosure of the clinical findings, to generate contextual information for interpreting their symptoms. These questions will be addressed to ascertain their relative importance to the patient and to assess the value of 'non-classical' indicators of a cancer diagnosis. The interviews will also help to identify optimal strategies for reassuring low risk patients that secondary referral is unnecessary, or efficient ways of discussing referral to secondary care with patients who present with a high probability of a cancer diagnosis. The interviews will also provide an opportunity to investigate patients' views about the clinical use of cancer decision instruments during consultations as means of decision-making about referral for secondary assessment. The interviews will examine the key factors, within the

context of the patients' presentation history, likely to trigger a GP consultation. Duration of symptoms is likely to impact on patient help seeking behaviour, and must therefore be explored in the context of the patient's decision to seek medical advice.

Interviews will be conducted mainly by one researcher in respondents' homes and last approximately 60 minutes. Two of the investigators (GL and TS) are qualitative researchers with a track record of conducting interviews with cancer patients and other vulnerable groups. They will provide training in interviewing cancer patients on sensitive subject matter, and supervise the researcher throughout the duration of the study. The initial interviews (approximately 4) will be conducted by one of the researchers, accompanied by a more junior researcher, to provide guidance on interview technique and discussion of sensitive subject matter with a vulnerable group of people. These initial interviews will be used to critically reflect on the interview content and discussion, and used to inform the conduct of subsequent interviews. Interviews will be structured and patient-led, so that patients will be encouraged to discuss topics that they feel comfortable with, and sensitive issues will only be broached if they have already been raised by the patient or if patients agree to discuss them. Interviews with both diagnostic groups will be conducted concurrently so that findings from one set of responses can be compared and contrasted with the other group, informing subsequent interviews and the topics to be addressed. An interview guide will ensure that the main questions are covered with each respondent, though issues specific to individual interviewees which are relevant to the research will also be discussed where appropriate. Data analysis will begin early on during fieldwork, and early insights generated from the interviews will be used in subsequent interviews. An iterative approach to data collection and analysis will be adopted, using inductive methods (e.g. thematic analysis). Due to the sensitive subject matter the consent procedure to the interviews will account for patients' changing health status and preferences at the time of the interview. In these circumstances we will accommodate their preferences by stopping or rescheduling the interview if patients express a desire to do so. The conduct of the interviews will also require a high degree of sensitivity, in which the qualitative team have extensive experience. The word 'cancer' or 'tumour' will not be used by the researcher during the interviews, even if the respondent has used the term.

**Analysis:** The interviews will be analysed by the researcher, under the supervision of the senior qualitative researchers (GL, LB and TS), using a variation of the constant comparative method derived from grounded theory. A thematic approach to the data analysis will be adopted (in contrast to narrative or case-based analysis), so that findings are compared within and between diagnostic groups. This will enable us to identify similarities and differences in non-classical symptoms across the two groups of patients. The analysis will involve a modified approach to grounded theory (Strauss and Corbin 1990), where themes derived from earlier analyses will feed into subsequent analyses, in a cycle of hypothesis generation and data interpretation. The interviews will be transcribed verbatim and coded using QSR's N-Vivo qualitative data management software. The Social Science Group at Keele University possess extensive expertise in the use of this software. The coded text will then be analysed in search of key themes. Each researcher will read and code the interview transcripts independently. Transcripts will then be compared for differences and similarities and in search of 'negative'

cases or ‘outliers’. The transcripts will be analysed to capture the onset, duration and fluctuation of symptoms and help seeking experiences of patients, so that each patients’ illness pathways are examined in context, paying particular attention to lay referral triggers, advice seeking, and lay decision-making prior to consultation with the GP. The interviews will therefore depict patients’ unique presentation history, rationale for seeking medical help, and symptom experience. Although the aim of the interviews is to detect non-classical symptomatic ‘indicators’ of a cancer diagnosis, they will also help to identify respondents’ reasoning behind their risk calculations; so that they are located in the context of their symptoms and presentation history rather than in isolation. Consequently, we will show which symptoms appear to be more salient and why.

Appendix 2.

## Delphi study

### Aims

To obtain consensus on the symptoms and signs to be included in the structured clinical examination of patients presenting with symptoms that could represent lung or colorectal cancer in primary care.

### Expert panel

A multidisciplinary panel of content and clinical experts will be invited, including clinicians and researchers in primary care and secondary care, both nationally and internationally. In order to obtain reliable results, a Delphi panel minimally needs to consist of at least 10 to 15 experts. More participants will add to the reliability, but will elaborate the procedure. We aimed to compose an expert panel of 20 members, a number which is commonly seen in consensus based research. Accounting for non-response, we will approach at least 40 clinicians and experts in the field of lung and colorectal cancer. All will be provided with an information letter explaining the aims and procedures of the Delphi study, and consent form.

### Round 1

The number of rounds may vary, but three rounds are expected to be sufficient. In the first round panel members will receive a long list of all potential diagnostic indicators derived from previously conducted systematic reviews and recent cohort studies and from the patient interviews. Separate lists will be compiled for symptoms and signs related to lung cancer and colorectal cancer. The panel members will be asked to score each diagnostic indicator for importance in the identification of lung or colorectal cancer on a 5 point likert scale (i.e. 1= not at all important, to 5= very important). Panel members are encouraged to propose additional diagnostic indicators that are not included in the list. The experts respond independently to the questions, and responses are confidential and anonymous.

The responses of all panel members will be collated and a reduced list of diagnostic indicators will be drawn up. A mean rating of 3 points will be considered to be an acceptable level of importance, and these variables are maintained for the next round. All diagnostic indicators will be ranked according to their rated importance. Newly suggested indicators will be added to the list and arranged by the frequency with which they are suggested.

### Round 2

The aim of the second round is to achieve consensus on the most important diagnostic indicators. The panel will receive feedback on the results of round 1, and will subsequently be asked to re-score all diagnostic indicators for importance. The panel will also be asked to rank

the 10 most important indicators by assigning scores between 10 (most important) to 1 (least important). Based on the total of scores from all panel members the diagnostic indicators will be ranked according to their argued diagnostic performance, including those newly suggested by panel members in the first round.

**Round 3**

In the third and final round panel members will be asked if they agree with the selection of diagnostic indicators resulting from the first two Delphi rounds. In case of disagreement panel members can alter the selection by replacing a maximum of 3 diagnostic indicators. Variables can be eliminated from the selection or be replaced by others. The 3rd Delphi round will also be used to ask the panel members' opinion regarding the amount of diagnostic uncertainty that would be tolerated, i.e. the maximum false positive and false negative rates that may be acceptable in a primary care population of patients with a first presentation of symptoms that could represent cancer. The final list of diagnostic indicators will be sent to panel members for final approval and to ask for feedback regarding the way of scoring the selected symptoms and signs, and issues surrounding user acceptability.

- *The use of a simple unweighted score may be clinically attractive but is probably a bad idea. The elegance of web-based data collection lends itself to computer-based calculator of probabilities that can handle more complex weighting, interactions, and non-linear functions without burdening the practitioner or patient. Concerns of over fitting the data will be addressed more directly by the validation sample although a strategy for how to update or revise the derived prediction model will be needed.*

This is an important point. We agree that computer systems may provide an opportunity to use weighted scores in clinical practice, and propose exploring this in the current study. However unweighted scores are clinically attractive and more likely to be used in practice - and it is probably no coincidence that probably the two most commonly used rules in primary care are the Ottowa ankle rules and the Centor criteria, and both are unweighted. Furthermore, unweighted scores are also statistically attractive: following Dawes' s classic observations there is growing empirical evidence that unweighted scores may perform better in validation exercises [36 37-39]. (The mechanism is unclear but may reflect a kind of 'shrinkage' in the statistical sense.) We propose to build both unweighted and weighted scores in the derivation samples, and then test both in the validation sample. We will proceed with the weighted score only if it outperforms the unweighted score in the validation sample.