Marcela Ewing, Peter Naredi, Chenyang Zhang and Jörgen Månsson

Identification of patients with non-metastatic colorectal cancer in primary care:

a case-control study

Abstract

Background
Colorectal cancer is the third most common cancer worldwide and second most common in Europe. Despite screening, it is often diagnosed at an unfavourable stage.

Aim
To identify and quantify features of non-metastatic colorectal cancer in primary care to enable earlier diagnosis by GPs.

Design and setting
A case-control study was conducted using diagnostic codes from national and regional healthcare databases in Sweden.

Method
A total of 542 patients diagnosed with non-metastatic colorectal cancer in 2011 and 2139 matched controls were selected from the Swedish Cancer Register (SCR) and a regional healthcare database respectively. All diagnostic codes (according to ICD-10) from primary care consultations registered the year before the date of cancer diagnosis (according to the SCR) were collected from the regional database. Odds ratios were calculated for variables independently associated with non-metastatic colorectal cancer using multivariable conditional logistic regressions. Positive predictive values (PPVs) of these variables were calculated, both individually and in combination with each other.

Results
Five features were associated with colorectal cancer before diagnosis: bleeding, including rectal bleeding, melaena, and gastrointestinal bleeding (PPV 3.9%, 95% confidence interval [CI] = 2.3 to 6.3); anaemia (PPV 1.4%, 95% CI = 1.1 to 1.8); change in bowel habit (PPV 1.1%, 95% CI = 0.9 to 1.5; abdominal pain (PPV 0.9%, 95% CI = 0.7 to 1.1); and weight loss (PPV 1.0%, 95% CI = 0.3 to 3.0); all *P*-value <0.05. The combination of bleeding and change in bowel habit had a PPV of 13.7% (95% CI = 2.1 to 54.4); for bleeding combined with abdominal pain this was 12.2% (95% CI = 1.8 to 51.2). A risk assessment tool for non-metastatic colorectal cancer was designed.

Conclusion
Bleeding combined with either diarrhoea, constipation, change in bowel habit, or abdominal pain are the most powerful predictors of non-metastatic colorectal cancer and should result in prompt referral for colorectal investigation.

Keywords
colorectal cancer; diagnosis; general practice; primary health care; Sweden.

INTRODUCTION

Patients diagnosed with non-metastatic colorectal cancer have a good survival outcome, but the risk of dying from metastatic colorectal cancer is high.[1,2] Colorectal cancer is the third most common cancer worldwide, with more than 1.3 million cases reported annually.[3] In Europe, it is the second most common cancer: more than 342 000 patients are diagnosed with colorectal cancer, and 150 000 die from it every year.[4] In Sweden, 6451 patients were diagnosed with colorectal cancer in 2014, and 2771 died from it.[5,6] Even though Sweden, from an international perspective, has high survival rates for many cancer diagnoses, colorectal cancer still has poor survival rates when discovered at an advanced stage;[7] late presentation and delays in diagnosis and treatment can be one cause for this.[8–13]

In Western countries such as Sweden, Norway, Denmark, and France, approximately 70–85% of patients with cancer are diagnosed in the primary care setting.[13-15] The evidence in literature is growing concerning the association between the length of the diagnostic interval and outcome;[11–13] as such, health practitioners must improve their knowledge about the features of colorectal cancer at early stages. Important contributions from research into early detection of colorectal cancer in primary care have been made in Norway, Denmark, and the UK.[16–19] The National Institute for Health and Care Excellence (NICE) published new guidelines for suspected cancer in 2015, and found evidence from 25–30 studies on single symptoms for colorectal cancer.[20] However, only nine studies reported on the cardinal symptom of rectal bleeding combined with other symptoms, and only two of these reported on other symptom combinations;[18,21] thus, there are only a few multi-symptom studies on colorectal cancer.

The colorectal cancer assessment tool for primary care and QCancer, which is a risk prediction algorithm based on both symptoms and risk factors, have both been developed for the UK. It is unlikely that primary care patients in Sweden have different symptom patterns from those of patients in the UK; nevertheless, the UK has poorer survival rates for colorectal cancer compared with Sweden and other European countries, which is thought to be partly related to late presentation and different stage distribution at diagnosis.[7,22,23] Due to these circumstances, and because the existing risk assessment tools have been developed for British conditions in a British primary care setting, it may not be appropriate to use them elsewhere. In addition, the British tool does not discriminate between non-metastasised and metastasised colorectal cancer.

Patients in Sweden with stage III colorectal cancer have good survival rates,

M Ewing, MD, PhD student; **J Månsson**, MD, PhD, professor of general medicine, Department of Public Health and Community Medicine/Primary Health Care; **P Naredi**, MD, PhD, professor of surgery Department of Surgery, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden. **C Zhang**, MSc, statistician, Regional Cancer Centre West, Sahlgrenska University Hospital, Gothenburg, Gothenburg, Sweden.

Address for correspondence
Marcela Ewing, University of Gothenburg, Institute of Medicine, Department of Primary Health Care, Box 454, 405 30 Göteborg, Sweden.

E-mail: marcela.ewing@rccvast.se

Submitted: 26 May 2016; **Editor's response:** 20 June 2016; **final acceptance:** 11 July 2016.

How this fits in

Rectal bleeding is recognised as one of the cardinal symptoms of colorectal cancer but has a low positive predictive value in a primary care setting. This study shows that bleeding combined with either diarrhoea, constipation, change in bowel habit, or abdominal pain are the most powerful predictors of non-metastatic colorectal cancer in primary care.

even when compared with patients who have stage II disease: figures from the Swedish colorectal cancer quality register for 2015 indicate that 5-year relative survival for patients with colon cancer who had undergone elective surgery was 93% for stage II and 76% for stage III.[1] For rectal cancer, regardless of mode of surgery (elective or non-elective), the 5-year survival for stage II was 86% and 71% for stage III.[2]

In order to increase survival rates for people with colorectal cancer, the most important factor is to be able to identify patients with a potentially curable disease. Sweden has high survival rates in colorectal cancer, despite not having a national screening programme. In contrast with both the UK and Denmark, however, primary care practitioners in Sweden do not have a role as 'gatekeepers'. These different conditions mean there is a need for risk assessment tools that can be used to detect colorectal cancer early in different primary care settings. As Sweden possesses unique total population-based databases, a case-control study was conducted using regional healthcare databases and a national cancer register. This study aimed to:

- identify and quantify the clinical features of non-metastatic colorectal cancer in primary care, both as single symptoms and in combinations; and
- develop a risk assessment tool for non-metastatic colorectal cancer for use in primary care.

METHOD

Study design

A total population-based, case-control study using the Swedish Cancer Register and a regional healthcare database in Region Västra Götaland (RVG), Sweden, was designed. This region, which has 1.6 million inhabitants (17% of the Swedish population), is situated in the south-west of the country and includes both rural and urban areas.

The Swedish Cancer Register, which was established in 1958, is one of the oldest disease registers in Sweden and has high validity.[24] All physicians, including pathologists, in Sweden are obliged by law to report all incident cases of cancer from both living and dead patients to the Swedish Cancer Register.[25] Each patient has a unique personal identity number, which all Swedish residents acquire either at birth or when they immigrate to Sweden.

The regional healthcare database was established in RVG in 2000. It covers all hospitals, specialised outpatient care, and all private and public primary healthcare centres. The database includes place of residence, age, sex, healthcare contacts, and diagnostic codes for diagnoses and surgical procedures.[26] Physicians are obliged to enter codes for a patient's current disease(s) or symptoms into the patient's medical records at each consultation. The reimbursement system for primary care providers is partly based on the disease burden of the patients, which is identified by diagnostic codes reported to this database.

Study population

All patients with colorectal cancer diagnosed in 2011 in RVG were identified from the Swedish Cancer Register. Patients and matched controls were investigated for primary care diagnostic profiles. Inclusion criteria were:

- diagnosed in RVG with colorectal cancer;
- alive at the time of the cancer diagnosis;
- aged ≥18 years; and
- visited the GP during the year before cancer diagnosis.

Individuals were excluded from participation if they:

- lacked controls;
- had a previous cancer diagnosis in the Swedish Cancer Register (1991–2010); or
- had a metastasised stage IV colorectal cancer.

Patients with previous cancer (that is, another cancer diagnosis registered in the Swedish Cancer Register during the 20-year period before 2011) were deliberately omitted to avoid consultations in primary care being a control or concern of previous cancer.

Controls were selected from the regional healthcare database. They had the same inclusion criteria as the patients with cancer, with the exception of the cancer diagnosis. Only controls from RVG who had visited a GP in primary care between

1 January 2010 and 31 December 2011 were eligible. Four controls were matched to each case on age, sex, and primary care unit.

Data collection and study measurements

The unique personal identity numbers of both cases and controls were linked to the regional healthcare database. All data concerning diagnoses and dates of consultations with a GP between 1 January 2010 and 31 December 2011 were collected. The data extracted included diagnostic codes according to the:

- Swedish version of the *International Statistical Classification of Diseases and Related Health Problems, 10th revision* (ICD-10); or
- *Classification of Diseases and Health Problems 1997 Primary Care* (KSH97-P) — this is an abbreviated version of ICD-10, adapted to Swedish primary care to facilitate diagnostic coding.[27–29]

Diagnostic codes

All the diagnostic codes registered when patients with cancer and their controls consulted their GP during the year preceding their cancer diagnoses were studied. As more than 6000 different diagnostic codes were received, their number was reduced according to clinical relevance. The codes were arranged by incidence in the study population, and the number reduced by merging the ICD-10 four-character diagnostic codes and KSH97-P codes to the closest three-character diagnostic code; once this was done, there were 575 diagnostic codes. The three-character codes are the 'core' classification and mandatory level for reporting to the World Health Organization's mortality database and for general international comparison.[30]

Data analysis

The 575 diagnostic codes for patients with non-metastatic colorectal cancer were used as variables for univariable conditional logistic regression. Those found to be associated with cancer entered multivariable analyses, after which a list of statistically significant variables associated with colorectal cancer was compiled.

A likelihood ratio was then empirically calculated for each variable (and combinations thereof); this ratio was calculated by dividing the probability of a patient with colorectal cancer being registered with a diagnostic code by the probability of a patient without cancer being registered with the same diagnostic code. Using the likelihood ratios, the incidence of the cancer diagnosis in question, and Bayes' theorem,[31] a positive predictive value (PPV) was calculated for each variable. Once there were PPVs for different symptoms and their combinations, a risk assessment instrument was created, which could be used as a clinical tool by GPs.

All analyses were performed using the statistical software R, (version 3.0.1).

RESULTS

Cases and controls

A flowchart of the study sample recruitment process is given in Figure 1. In total, 753 patients with colorectal cancer were identified in the Swedish Cancer Register. As the study focused on early features of non-metastatic cancer, 65 patients who had no stage noted or whose stage could not be classified as non-metastasised or metastasised colorectal cancer were omitted. Of the remaining 688 patients, 542 had stage I–III (78.8%) and 146 (21.2%) had stage IV cancer (metastatic cancer). All patients with stage IV cancer were excluded from the study, resulting in a final sample of 542 patients with non-metastatic cancer.

A total of 2152 controls were generated, but 13 died before diagnosis of their case. Included in the study were 2139 controls matched to patients with stage

Figure 1. Sample recruitment flowchart.

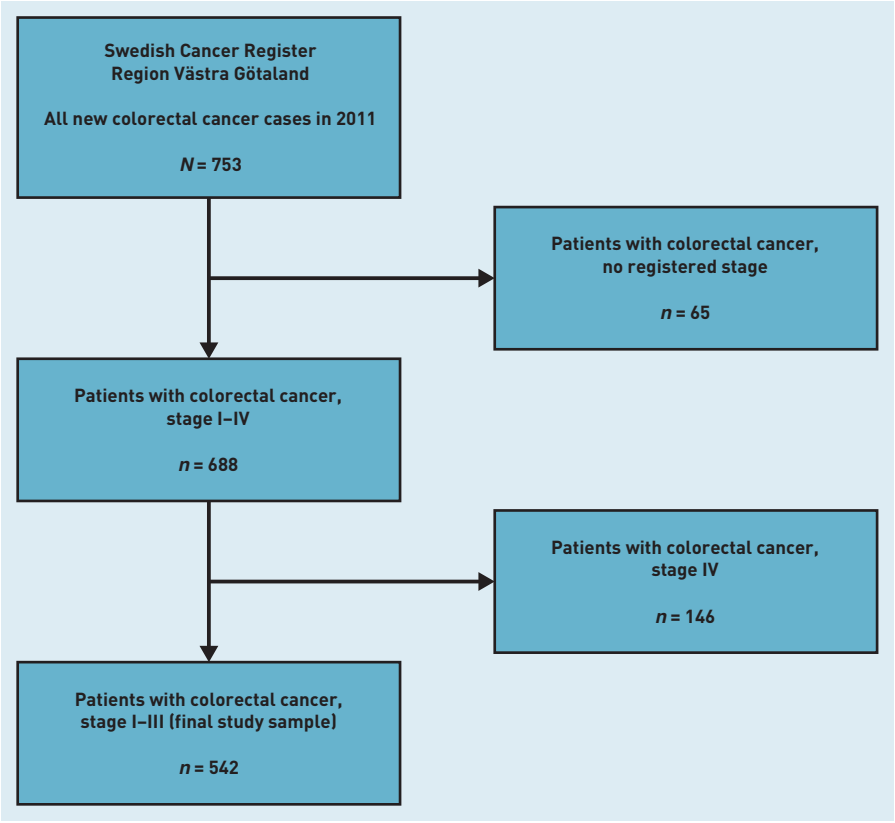

Table 1. Sample characteristics

Characteristic	Patients with colorectal cancer, *n* = 542	Controls, *n* = 2139
Median age at diagnosis, years (range)	72 (30–94)	72 (30–94)
<50 years, *n* (%)	31 (5.7)	116 (5.4)
50–59 years, *n* (%)	52 (9.6)	210 (9.8)
<60 years, *n* (%)	83 (15.3)	326 (15.2)
60–80 years, *n* (%)	347 (64.0)	1378 (64.4)
>80 years, *n* (%)	112 (20.7)	435 (20.3)
Stage I, *n* (%)	118 (21.8)	—
Stage II, *n* (%)	223 (41.1)	—
Stage III, *n* (%)	201 (37.1)	—
Median number of consultations per patient in year before cancer diagnosis, *n* (IQR)	5 (3–8)	4 (2–7)
Median number of unique diagnostic codes per patient in year before cancer diagnosis, *n* (IQR)	6 (4–10)	5 (3–8)

IQR = interquartile range.

I–III colorectal cancer. Table 1 outlines the characteristics of the study sample.

Variables

Table 2. Univariable analysis of diagnoses with odds ratio >1.5 in patients in primary care 12 months before colorectal cancer diagnosis

ICD-10 code and diagnosis	Prevalence, %	OR	*P*-value
K625 Haemorrhage of anus and rectum	3.7	79	<0.001
R194 Change in bowel habit	7.5	23	<0.001
K921 Melaena	4.6	19	<0.001
R634 Abnormal weight loss	0.7	13	0.023
R190 Intra-abdominal and pelvic swelling, mass, and lump	1.3	13	0.002
K922 Gastrointestinal haemorrhage	3.5	12	<0.001
D50 Iron deficiency anaemia	12.3	10	<0.001
D64 Other anaemias	17.0	6.5	<0.001
R63 Symptoms and signs concerning fluid and intake	1.1	4.5	0.012
R10 Abdominal and pelvic pain	22.4	4	<0.001
R14 Flatulence	1.5	3.5	0.009
K591 Functional diarrhoea	1.8	3.5	0.004
R58 Haemorrhage, not elsewhere classified	1.1	3.4	0.027
R11 Nausea and vomiting	2.0	2.7	0.011
K590 Constipation	7.5	2.6	<0.001

ICD-10 = International Statistical Classification of Diseases and Related Health Problems, 10th Revision.
OR = odds ratio.

All 575 variables occurred in ≥1% of either cases or controls. After univariable conditional logistic regression, 45 statistically significant variables (*P*-value threshold <0.05) were considered for multivariable analyses. As clinical features such as symptoms and signs were of interest, diagnostic codes for unspecific medical conditions or diseases that had no connection to colorectal cancer were excluded. Only variables with an odds ratio of >1.5 were retained, which left 15 variables (Table 2).

As several of the diagnostic codes represented similar clinical features, they were merged into 10 clinical groups. Variable bleeding included colorectal bleeding, melaena, gastrointestinal (GI) bleeding, and unclassified bleeding; variable loss of weight contained codes for both abnormal weight loss and anorexia; and variable anaemia was a fusion of iron deficiency anaemia and other anaemias. As it was not possible to classify the diagnostic code 'change in bowel habit' as either diarrhoea or constipation, these three variables were merged into one and classified as 'change in bowel habit'. This also helped ensure the tool created for use in primary care was simple, as it had fewer variables.

After these two selective processes, multivariable analysis was undertaken of the 10 variables. This showed that five symptoms or signs (*P*-value <0.05, although for most this was <0.001) were independently associated with colorectal cancer during the year preceding the cancer diagnosis. These were:

- change in bowel habit;
- bleeding;
- weight loss;
- abdominal pain; and
- anaemia.

Positive predictive values

Figure 2 shows the PPVs for non-metastatic colorectal cancer for those symptoms associated with the disease:

- independently in the multivariable analysis;
- in combination with another symptom;
- when the symptom was reported a second time;
- for patients aged ≥50 years (as they accounted for 94.3% of the patients with colorectal cancer); and
- against a background risk of 0.25%.

A combination of bleeding and a change

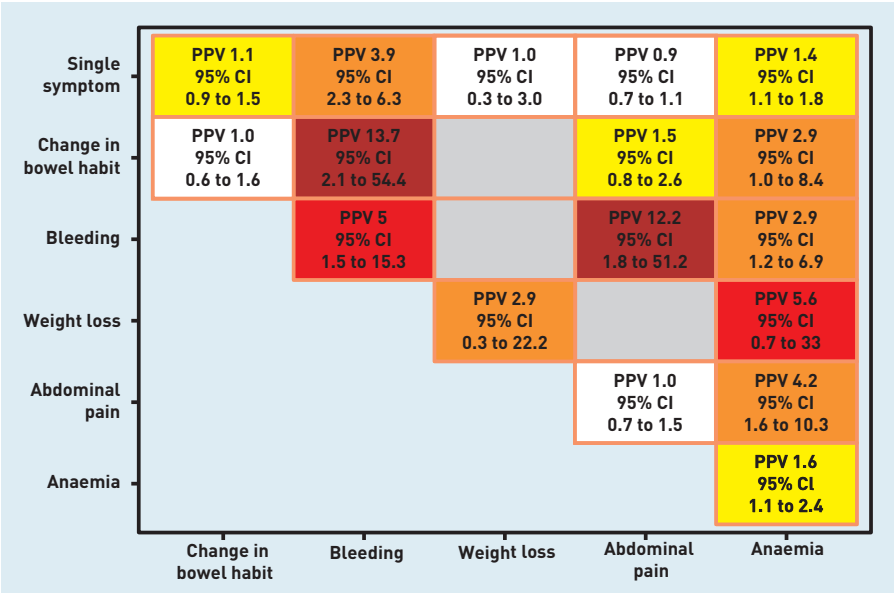

Figure 2. Risk assessment tool for non-metastatic colorectal cancer. Risk plot with PPV for colorectal cancer stage I–III, in patients aged ≥50 years (against a background risk of 0.25%). Top-row single symptoms show the individual risk of each symptom. The diagonal rows show the PPV when the symptom is reported a second time. Other cells show the PPV of the combination of two different symptoms. PPV = positive predictive value. White: 0–1%. Yellow: >1%. Orange: >2.5%. Red: >5%. Dark red: >10%. Grey: too few patients with this combination.

in bowel habit, or bleeding and abdominal pain, were found to be the combinations of clinical features with the highest PPVs of non-metastatic colorectal cancer.

DISCUSSION

Summary

Five symptoms and signs were found to be independently associated with colorectal cancer, and — more importantly — combinations of symptoms and signs with a high predictive value for non-metastatic colorectal cancer were identified. The combination of bleeding and a change in bowel habit, or bleeding and abdominal pain, had the highest PPVs of non-metastatic colorectal cancer. These findings enabled a risk assessment instrument for primary care to be created.

Strengths and limitations

The study is total population based, which is its main strength. All patients with cancer in the sample were identified via the Swedish Cancer Register, so there is no selection bias and the completeness of the register is very high.[24] Early features of colorectal cancer were studied as these have great implications in primary care. The symptoms and signs with which the risk assessment instrument was constructed are intended to help GPs detect non-metastatic colorectal cancer early. In addition, the instrument is simple to use without too many variables needing to be entered.

Another strength is the use of diagnostic codes; however, this could also be considered a limitation as not all the symptoms for which a patient consults their GP will be recorded as a diagnostic code in their medical record. Important information about symptoms hidden in the free text of the medical record can be lost, which has been observed in other fields of research in primary care databases such as that of rheumatoid arthritis.[32] In addition, for the purposes of this study diagnostic codes were merged, from four- to three-character diagnostic codes. These were finally merged into 10 clinical groups, which explains why GI bleeding from both the upper and lower GI tract were merged into variable bleeding. This could be considered a limitation too.

The time span studied may also be considered a limitation. Even though many studies suggest that most cancer symptoms occur 3–6 months before the cancer diagnosis,[18,33,34] a longer time span than the one in this study may be needed for observation. In addition, some cells in the risk assessment tool lack a PPV because there were too few cases — for example, as with weight loss. It could be that Swedish GPs are underdiagnosing and tend to use disease diagnostic codes rather than those that are symptom based. Finally, it is important to be aware that the risk assessment tool that was developed has not yet been validated.

Comparison with existing literature

The risk assessment tool for colorectal cancer that was developed for primary care in the UK has PPVs for single symptoms and also for pairs of symptoms that are similar to those outlined in the tool created here.[18,35] However, in the results presented here, diarrhoea, constipation, and change in bowel habit were merged into a single variable (change in bowel habit). The combination of bleeding and change in bowel habit yielded a PPV of 13.7% (95% CI = 2.1 to 54.4); this is four times higher than PPVs of the combination of rectal bleeding and constipation or diarrhoea in the UK risk assessment tool. In addition, this study showed that bleeding combined with abdominal pain had a PPV of 12.2% (95% CI = 1.8 to 51.2), which is almost four times higher than in the UK risk assessment tool.[18] Most importantly, however, the study outlined here showed that these risk combinations are presented in patients with non-metastatic colorectal cancer.

In a systematic review from 2011 — the aim of which was to investigate the diagnostic value of symptoms for colorectal cancer in primary care — the summary estimate PPV of rectal bleeding and change

in bowel habit was 11.8%, which is similar to the findings presented here; however, that review did not study only non-metastatic colorectal cancer.[36] Another systematic review from primary care settings of only early-stage colorectal cancer found PPVs of 9–12% with these combined symptoms.[37]

QCancer is a risk prediction algorithm developed in the UK to identify an individual's absolute risk of having a diagnosis of colorectal cancer in the next 2 years, based on both symptoms and risk factors. When compared with specific symptoms such as rectal bleeding, abdominal pain, weight loss, and anaemia alone, the PPVs from the study presented here are very similar.[22] The QCancer algorithm is based on colorectal cancer diagnoses but makes no distinction between early or advanced colorectal cancer. The risk assessment tool created in this study has been designed to diagnose colorectal cancer at an earlier, more favourable stage than other risk assessment instruments.

A recently-published study from Israel describes a validated computational model to identify individuals at increased risk of colorectal cancer at an earlier stage by analysing complete blood counts, age, and sex.[38] This model is not symptom based and can detect colorectal cancers in asymptomatic and even non-anaemic patients; this adds a significant clinical benefit to symptom-based tools but, as the model is based on changes in haemoglobin over time, detection depends on continuity in consultation pattern, regular complete blood counts, and utilisation of the same analysing laboratory facilities. The advantage of the symptom-based tool developed in this study, compared with the model above, is that it is more flexible and not dependent on patients needing to consult the same primary care unit or laboratory for blood samples.

Implications for research and practice

A useful and simple risk assessment tool has been developed for GPs to use in everyday practice. Even though this tool has not yet been validated it could help GPs to detect colorectal cancer at an early stage. A change in bowel habit — whether constipation, diarrhoea, or any change in bowel habit — in combination with bleeding, as well as bleeding combined with abdominal pain, should cause GPs to consider a swift referral to confirm or exclude colorectal cancer.

These two symptom combinations had the highest PPVs in the study presented here but, as recent NICE guidelines applied a new 3% threshold for the PPV warranting urgent referral,[20] the research presented here indicates that there are several other pairs of symptoms that should lead to prompt clinical action, such as early referral for colorectal investigation. As the majority of patients present to primary care, GPs could use this tool to help identify those at high risk of non-metastatic colorectal cancer.

There are many potential benefits, as well as challenges, about the use of risk prediction tools for cancer in primary care and their implementations. Further validation of different risk prediction tools to assess the acceptability, clinical impact, and economic implications are needed.[39]

Funding

The study was conducted without external funding. Access to the regional healthcare database, VEGA, was financed by Regional Cancer Centre West, Sahlgrenska University Hospital, Gothenburg, Sweden.

Ethical approval

The Regional Ethical Review Board in Gothenburg approved the study protocol (reference number: 252–12).

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Acknowledgements

The authors thank Erik Holmberg, statistician at Regional Cancer Centre West, for the extraction of data from the Swedish Cancer Register, as well as Kristina Narbro and Mona-Lis Dalbrekt from Västra Götaland Region's Department of Health Care Evaluation for their help in extracting data from VEGA, the regional healthcare database.

Discuss this article

Contribute and read comments about this article: **bjgp.org/letters**

REFERENCES

1. Regionala Cancercentrum. *Koloncancer: nationell kvalitetsrapport för år 2015 från Svenska Kolorektalcancerregistret* [Colon cancer: national quality report for year 2015 from the Swedish Colorectal Cancer Register]. http://www.cancercentrum.se/globalassets/cancerdiagnoser/tjock--och-andtarm/kvalitetsregister/nationell-kvalitetsregisterrapport-tjocktarmscancer-2015.pdf (accessed 20 Sep 2016).
2. Regionala Cancercentrum. *Rektalcancer: nationell kvalitetsrapport för år 2015 från Svenska Kolorektalcancerregistret.* [Rectal cancer: national quality report for year 2015 from the Swedish Colorectal Cancer Register]. http://www.cancercentrum.se/globalassets/cancerdiagnoser/tjock--och-andtarm/kvalitetsregister/nationell-kvalitetsregisterrapport-andtarmscancer-2015.pdf (accessed 20 Sep 2016).
3. World Health Organization. International Agency For Research on Cancer. *GLOBOCAN 2012: estimated cancer incidence, mortality and prevalence worldwide in 2012.* http://globocan.iarc.fr/Pages/fact_sheets_population.aspx (accessed 20 Sep 2016).
4. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, *et al.* Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* 2013; **49(6)**: 1374–1403.
5. Socialstyrelsen. *Cancerincidens i Sverige 2014.* [Cancer incidence in Sweden 2014]. Stockholm: Socialstyrelsen, 2015. https://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/20008/2015-12-26.pdf (accessed 29 Sep 2016).
6. Socialstyrelsen. *Dödsorsaker 2014.* [Causes of death 2014]. http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/19909/2015-8-1.pdf (accessed 20 Sep 2016).
7. De Angelis R, Sant M, Coleman MP, *et al.* Cancer survival in Europe 1999–2007 by country and age: results of EUROCARE-5 — a population-based study. *Lancet Oncol* 2014; **15(1)**: 23–34.
8. Lyratzopoulos G, Wardle J, Rubin G. Rethinking diagnostic delay in cancer: how difficult is the diagnosis? *BMJ* 2014; **349:** g7400.
9. Macleod U, Mitchell ED, Burgess C, *et al.* Risk factors for delayed presentation and referral of symptomatic cancer: evidence for common cancers. *Br J Cancer* 2009; **101(Suppl 2):** S92–S101.
10. Langenbach MR, Sauerland S, Kröbel KW, Zirngibl H. Why so late?! — Delay in treatment of colorectal cancer is socially determined. *Langenbecks Arch Surg* 2010; **395(8)**: 1017–1024.
11. Neal RD, Tharmanathan P, France B, *et al.* Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? Systematic review. *Br J Cancer* 2015; **112(Suppl1)**: S92–S107.
12. Torring ML, Frydenberg M, Hamilton W, *et al.* Diagnostic interval and mortality in colorectal cancer: U-shaped association demonstrated for three different datasets. *J Clin Epidemiol* 2012; **65(6):** 669–678.
13. Torring ML, Frydenberg M, Hansen RP, *et al.* Evidence of increasing mortality with longer diagnostic intervals for five common cancers: a cohort study in primary care. *Eur J Cancer* 2013; **49(9):** 2187–2198.
14. Månsson J. *The diagnostic process of cancer from the general practitioner's point of view.* PhD thesis. Göteborg: Göteborgs Universitet, 1999. https://gupea.ub.gu.se/handle/2077/13126 (accessed 20 Oct 2016).
15. Demagny L, Holtedahl K, Bachimont J, *et al.* General practitioners' role in cancer care: a French-Norwegian study. *BMC Res Notes* 2009; **2(1)**: 200.
16. Holtedahl K. *Early diagnosis of cancer in general practice: a manual.* Berlin Heidelberg: Springer-Verlag, 1990.
17. Hansen PL, Hjertholm P, Vedsted P. Increased diagnostic activity in general practice during the year preceding colorectal cancer diagnosis. *Int J Cancer* 2015; **137(3):** 615–624.
18. Hamilton W, Round A, Sharp D, Peters TJ. Clinical features of colorectal cancer before diagnosis: a population-based case-control study. *Br J Cancer* 2005; **93(4):** 399–405.
19. Hamilton W, Lancashire R, Sharp D, *et al.* The risk of colorectal cancer with symptoms at different ages and between the sexes: a case–control study. *BMC Med* 2009; **7:** 17.
20. National Institute for Health and Care Excellence. *Suspected cancer: recognition and referral.* NG12. London: NICE, 2015. https://www.nice.org.uk/guidance/NG12 (accessed 30 Sep 2016).
21. Meineche-Schmidt V, Jorgensen T. 'Alarm symptoms' in patients with dyspepsia: a three-year prospective study from general practice. *Scand J Gastroenterol* 2002; **37(9):** 999–1007.
22. Hippisley-Cox J, Coupland C. Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2012; DOI: 10.3399/bjgp12X616346.
23. Maringe C, Walters S, Rachet B, *et al.* Stage at diagnosis and colorectal cancer survival in six high-income countries: a population-based study of patients diagnosed during 2000–2007. *Acta Oncol* 2013; **52(5):** 919–932.
24. Barlow L, Westergren K, Holmberg L, Talbäck M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. *Acta Oncol* 2009; **48(1):** 27–33.
25. Socialstyrelsen. *Cancerförekomst i Sverige 2011.* [Cancer incidence in Sweden 2011]. Stockholm: Socialstyrelsen, 2012. http://www.socialstyrelsen.se/publikationer2012/2012-12-19 (accessed 29 Sep 2016).
26. Björck S, Palaszewski B, Friberg L, Bergfeldt L. Atrial fibrillation, stroke risk, and warfarin therapy revisited: a population-based study. *Stroke* 2013; **44(11):** 3103–3108.
27. Socialstyrelsen. *Internationell statistisk klassifikation av sjukdomar och relaterade hälsoproblem-systematisk förteckning (ICD-10-SE).* [International statistical classification of diseases and related Health Problems, 10th revision (ICD-10)]. Stockholm: Socialstyrelsen. 2010.
28. Socialstyrelsen. *Klassifikation av sjukdomar och hälsoproblem 1997. Primärvård.* KSH97-P. [Classification of diseases and health problems 1997. Primary care]. Stockholm: Socialstyrelsen, 1997.
29. Nyström M, Vikström A, Nilsson GH, *et al.* Enriching a primary health care version of ICD-10 using SNOMED CT mapping. *J Biomed Semantics* 2010; **1(1):** 7.
30. World Health Organization. *International statistical classification of diseases and related health problems, 10th revision, volume 2. Instruction manual.* Geneva: WHO, 2010.
31. Papoulis A. 'Bayes' theorem in statistics' and 'Bayes' theorem in statistics (reexamined)'. In: *Probability, random variables, and stochastic processes.* 2nd edn. New York, NY: McGraw-Hill, 1984: 38–114.
32. Ford E, Nicholson A, Koeling R, *et al.* Optimising the use of electronic health records to estimate the incidence of rheumatoid arthritis in primary care: what information is hidden in free text? *BMC Med Res Methodol* 2013; **13:** 105.
33. Biswas M, Ades AE, Hamilton W. Symptom lead times in lung and colorectal cancers: what are the benefits of symptom-based approaches to early diagnosis? *Br J Cancer* 2015; **112(2)**: 271–277.
34. Hansen PL, Hjertholm P, Vedsted P. Increased diagnostic activity in general practice during the year preceding colorectal cancer diagnosis. *Int J Cancer* 2015; **137(3):** 615–624.
35. Hamilton W, Green T, Martins T, *et al.* Evaluation of risk assessment tools for suspected cancer in general practice: a cohort study. *Br J Gen Pract* 2013; DOI: 10.3399/bjgp13X660751.
36. Astin M, Griffin T, Neal RD, *et al.* The diagnostic value of symptoms for colorectal cancer in primary care: a systematic review. *Br J Gen Pract* 2011; DOI: 10.3399/bjgp11X572427.
37. Shapley M, Mansell G, Jordan JL, Jordan KP. Positive predictive values of ≥5% in primary care for cancer: systematic review. *Br J Gen Pract* 2010; DOI: 10.3399/bjgp10X515412.
38. Kinar Y, Kalkstein N, Akiva P, *et al.* Development and validation of a predictive model for detection of colorectal cancer in primary care by analysis of complete blood counts: a binational retrospective study. *J Am Med Inform Assoc* 2016; **23(5):** 879–890.
39. Usher-Smith J, Emery J, Hamilton W, *et al.* Risk prediction tools for cancer in primary care. *Br J Cancer.* 2015; **113(12):** 1645–1650.