Identification and early detection of cancer patients in primary care

Marcela Ewing

Department of Public Health and Community Medicine
Institute of Medicine
Sahlgrenska Academy at University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2018

Cover illustration: Be different. www.istockphoto.com

Identification and early detection of cancer patients in primary care

marcela.ewing@rccvast.se

ISBN 978-91-629-0402-9 (PRINT); 978-91-629-0403-6 (PDF)
http://hdl.handle.net/2077/55379

Printed by BrandFactory in Gothenburg, Sweden 2018

To two fantastic women
my mother and daughter

Identification and early detection of cancer patients in primary care

Marcela Ewing

Department of Public Health and Community Medicine, Institute of Medicine
Sahlgrenska Academy at University of Gothenburg
Gothenburg, Sweden

ABSTRACT

Aim The aim of this thesis was to investigate how general practitioners (GP) can identify patients in primary care with potential common cancers, at an early stage. It was also to design a risk assessment tool for colorectal cancer.

Method Four population-based case-control studies were conducted with cancer patients diagnosed in 2011 in Region Västra Götaland, Sweden, with prostate, breast, colorectal, lung, gynaecological, and skin cancers, including malignant melanoma. Data were retrieved from the Swedish Cancer Register, the regional healthcare database and the regional repository for radiology.

Results The patients' frequency of consultation in primary care increased 50–100 days before cancer diagnosis (Paper I). More than half had consulted a GP at least four times in the year before cancer diagnosis. A considerable proportion of patients presented with early clinical features that were focal and had benign characteristics (Paper II). Bleeding combined with diarrhoea, constipation, a change in bowel habit, or abdominal pain had the highest positive predictive values of non-metastatic colorectal cancer. A risk assessment tool was designed for colorectal cancer (Paper III). Non-metastatic lung cancer could not be identified by clinical features (Paper IV).

Conclusion Increased consultation frequency in primary care is a risk marker for common cancers as are focal features presented with benign characteristics. It is possible for a GP to identify patients with non-metastatic colorectal cancer by their clinical features. There is not enough evidence to suggest that patients with non-metastatic lung cancer can be identified.

Keywords: cancer; consultation; diagnosis; early detection; general practice; primary health care.

ISBN: 978-91-629-0402-9 (PRINT); 978-91-629-0403-6 (PDF)

SAMMANFATTNING PÅ SVENSKA

Bakgrund Cancersjukdomar är en vanlig orsak till sjukdom och död, både globalt och i Sverige. I Sverige insjuknar cirka 61 000 personer årligen i någon form av cancer. Patienterna söker oftast för sina besvär i primärvården. Allmänläkarna är därför de som oftast påbörjar utredning av patienter där symtom eller fynd väcker cancermisstanke och senare resulterar i en cancerdiagnos. Syftet med avhandlingen var att ta reda på hur allmänläkare kan känna igen patienter som har tecken på någon av de vanligaste cancersjukdomarna, om det är möjligt att upptäcka cancer i ett tidigt skede samt att utarbeta ett riskvärderingsinstrument för tjock- och ändtarmscancer.

Metod Fyra fall-kontrollstudier med sammanlagt 4562 cancerpatienter och 17 979 kontrollpatienter utan cancer genomfördes. Vi samlade in uppgifter om alla vuxna patienter i Västra Götalandsregionen som under 2011 diagnosticerades med prostata-, bröst-, tjock- och ändtarm-, lung-, gynekologisk eller hudcancer inklusive malignt melanom. Uppgifter om cancerdiagnoser, diagnosdatum, tumörstadium, diagnoskoder samt innehåll i remisser till lungröntgen hämtades från cancerregistret, regionala hälsodatabasen VEGA samt det regionala bild- och funktionsregistret.

Resultat Patienterna som senare fick en cancerdiagnos började söka läkare i primärvården mer frekvent 50–100 dagar före sin diagnos (delarbete I). Mer än hälften av patienterna besökte allmänläkare fyra eller fler gånger året innan de fick sin cancerdiagnos. Av dem som sökte läkare ofta men utan tydliga varningstecken på cancer sökte många redan de två första gångerna med symtom som visade sig vara associerade med cancer. Dessa symtom kom från en bestämd del av kroppen och tedde sig godartade (delarbete II). Blödning från tarmen kombinerad med diarré, förstoppning, ändrade avföringsvanor eller smärta i buken var de symptom som var starkast förknippade med icke spridd tjock- och ändtarmscancer. Ett riskvärderingsinstrument utarbetades för denna cancer (delarbete III). Icke spridd lungcancer kunde ej identifieras utifrån symtom (delarbete IV).

Slutsats Det finns sätt för allmänläkare att urskilja patienter med misstänkt vanlig cancer. Ett varningstecken är när patienter plötsligt söker gång på gång i primärvården även med symptom som ter sig godartade. Detta kan vara ett tecken på en bakomliggande cancersjukdom. Det finns olika kombinationer av symtom från mage och tarm som gör att tidig tjock-och ändtarmscancer är möjlig att diagnosticera. Utifrån vår studie kunde vi ej säga att patienter med icke spridd lungcancer kunde kännas igen utifrån typen av symtom.

LIST OF PAPERS

This thesis is based on the following studies, referred to in the text by their Roman numerals.

I. Ewing M, Naredi P, Nemes S, Zhang C and Månsson J. Increased consultation frequency in primary care, a risk marker for cancer: a case–control study. *Scand J Prim Health Care*. 2016; **34(2)**: 2015-2212.

II. Ewing M, Naredi P, Zhang C and Månsson J. Diagnostic profile characteristics of cancer patients with frequent consultations in primary care before diagnosis: a case-control study. *Accepted for publication 8 Feb 2018 in Family Practice*.

III. Ewing M, Naredi P, Zhang C and Månsson J. Identification of patients with non-metastatic colorectal cancer in primary care: a case-control study. *Br J Gen Pract*. 2016; **66(653)**: e880-e886.

IV. Ewing M, Naredi P, Zhang C, Lindsköld L and Månsson J. Clinical features of lung cancer patients with non-metastatic disease in primary care: a case-control study. *BJGP Open*. 2018.

CONTENT

ABBREVIATIONS

COPD	Chronic obstructive pulmonary disease
CRC	Colorectal cancer
CPP	Cancer patient pathway
EIA	Enterprise information archive for radiology
GP	General practitioner
LC	Lung cancer
LDCT	Low-dose computed tomography
LR	Likelihood ratio
NICE	National Institute for Health and Care Excellence
OR	Odds ratio
PPV	Positive predictive value
RVG	Region Västra Götaland
SCR	Swedish Cancer Register
VEGA	The regional healthcare database
WHO	World Health Organisation

1 INTRODUCTION

‘How was it possible not to see that this patient suffered from cancer? The symptoms were typical of the disease. Don’t GPs examine their patients?’ A skilled oncologist colleague asked these questions (that shocked me a bit), while working as an oncologist at the University Hospital.

The challenge for an oncologist is to treat, and if possible cure patients with cancer that has been diagnosed by someone else. But was it really so simple to diagnose cancer, which can mean any of the two hundred diseases that share this name?

When I changed my clinical path from oncology to primary health care and was consulted every day by several patients with symptoms and signs of which cancer was one of the differential diagnoses, my clinical experience taught me the answer. Yes, despite doing a thorough examination and investigation of the patient, you can easily miss typical cancer features because they are similar to the features of common, less serious diseases. At the time, I did not have any deeper academic knowledge in the field that could underpin my answer.

An interest in knowing more led me to the topic of this thesis. I wondered how a GP could recognize patients with symptoms and findings with high risk of having cancer. I asked myself, how can GPs select the right patient for the appropriate diagnostic investigation to confirm or exclude a cancer diagnosis from among the many patients that consult them for tiredness or cough or any other symptom or finding? And is it possible to detect cancer earlier, at a less advanced stage than is often the case when it comes to our most common cancers?

2 BACKGROUND

'Cancer' is a generic term for a large group of diseases that can affect any part of the body. The main characteristic of cancer is the creation of abnormal cells that grow beyond their normal boundaries and have the ability to spread or metastasize to other organs, which is the major cause of death from cancer.

A cancer diagnosis used to be perceived as a death sentence, but today a great proportion of patients that are treated, are either cured or live for many years with the disease. However, no matter how sophisticated the diagnostics or treatment modalities, a high mortality rate in many common cancers is due primarily to late-stage diagnosis and delay in treatment. There has been a lack of consensus over whether delays in cancer diagnosis truly affect survival.[1-3] However, an increasing number of studies have confirmed that screening and a timely diagnosis are associated with better clinical outcomes.[4-9]

2.1 Cancer epidemiology

Cancer is one of the leading causes of morbidity and mortality worldwide, with more than 14 million new cases and more than 8 million deaths in 2012.[10] The most commonly diagnosed cancers were lung, breast and colorectal; the cancers that most commonly caused death were lung, liver and stomach. In 2012 in Europe, there were 3.45 million new cases of cancer (excluding non-melanoma skin cancer) and 1.75 million deaths from cancer.[11] The most common cancers were cancers of the breast, colorectum, prostate and lung, that represented half of the cancer burden in Europe, with lung cancer as the most common cause of death from cancer.

In 2015, approximately 65,000 new cases of cancer for 61,000 individuals were reported to the Swedish Cancer Registry. The most common cancers were cancers of the prostate-, breast-, and skin.[12] Approximately a quarter of all deaths in Sweden is attributed to cancer.[13] Lung cancer is the most common cause of cancer-related death among women, and the next most common (after prostate cancer) for men.[14] This thesis is based on data registered in 2011 in the Swedish Cancer Registry. The cancer incidence was

57,726 cases in 46,286 persons in whom the cancer was diagnosed for the first time, 52% in men and 48% in women.[15]

2.2 Cancer detection

Cancer can be detected mainly in three ways; by screening, through symptoms or signs presented by patients, or simply by chance when the patient is being investigated for some other concern.

Many countries have implemented screening, in which they offer diagnostic testing to a target population at risk of developing certain cancers, to detect the cancer at an early stage when it is symptom-less. At present, Sweden does screening for breast and cervical cancer, and has plans to do colorectal cancer screening.[16] Although screening programmes have been shown to reduce mortality [4, 17, 18], they diagnose only a small part of all cancer patients. The majority of patients diagnosed with cancer present with symptoms in primary care.[19-21] In Western countries such as Sweden, Norway, Denmark and France, GPs are involved in initiating the diagnostic pathway in 70%–87% of patients later diagnosed with cancer. [7, 22-25]

2.3 Cancer survival and stage

Because each type of cancer has different biological profiles, the survival of people with cancers varies substantially, depending on two main prognostic factors. One is the differentiation of cancer cells, which is described as the cancer's aggressiveness.

Another crucial characteristic is the stage of the tumour, which is a major determinant of treatment and prognosis. Stage is determined by the TNM system, which describes the anatomical extent of disease, and is based on the assessment of three components. Thus, the stage is defined by the size of the tumour (T), the absence or presence of regional lymph nodes (N) and the absence or presence of distant metastasis (M). Different cancers have different classifications of stages, but generally the survival depends on whether the cancer is small and localized (Stage I), more advanced but still localized to one organ (Stage II) has invaded regional lymph nodes (Stage III) or has spread to other organs (Stage IV).[26]

A timely cancer diagnosis is important for the patient regardless of stage, but the outcome often depends on its stage at diagnosis. Survival is higher in cancers detected at an early stage because of screening or alarm features like a lump in breast. The 5-year survival rate observed in Sweden for breast cancer is 83%.[27] However, cancers such as lung cancer which are neither part of screening programmes nor have key features that are easily recognized are often diagnosed late which results in a poor prognosis.[7, 9, 28] Sweden has high survival rates for many types of cancer, but has poor survival rates for lung cancer.[29] The relative 5-year survival rate in lung cancer is 18%,[30] and half of the patients are diagnosed at Stage IV.[31] This stage distribution is similar in other countries; in the UK, half of lung cancer patients with a known stage were diagnosed at Stage IV.[32] The poor survival rate is thus mainly due to late stage diagnosis (Figure 1).

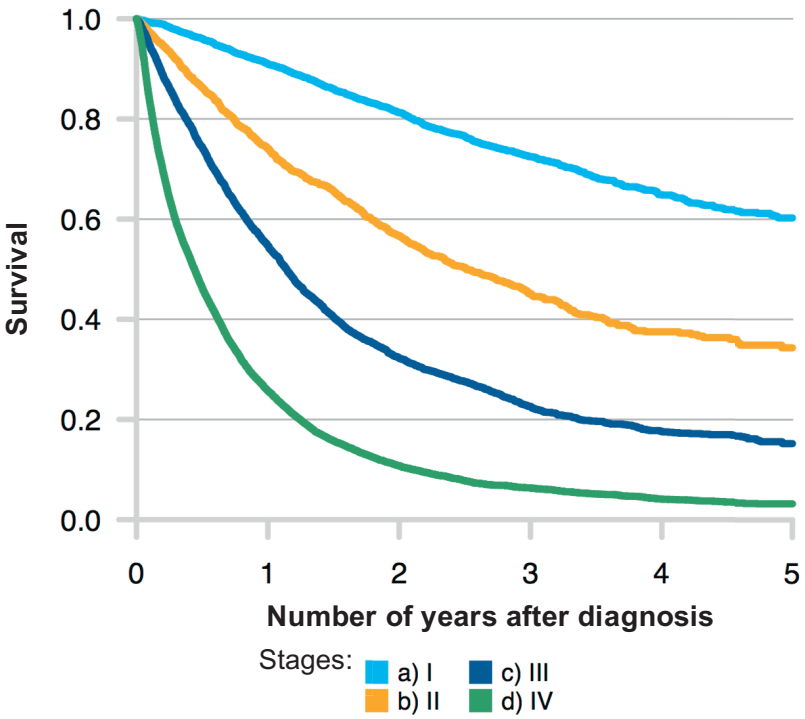

Figure 1. Survival in lung cancer in Sweden depending on stage at diagnosis in 2012-2016. Reprinted with permission from the Swedish Lung Cancer Register.[31]

Because lung cancer and colorectal cancer are two of the most common cancers worldwide and in Sweden, with high mortality especially in lung cancer, a timely cancer diagnosis should be highly prioritized. That is why these two cancers are highlighted in the thesis.

2.4 Cancer in primary care

In many countries, including Sweden, primary health care is the cornerstone of all other health services. GPs, who are specialists in general medicine, have a broad knowledge of and competence in the field of diseases and many kinds of medical conditions for the entire population. Because GPs in Sweden do not have a gate keeping function, as in some other European countries, patients are able to consult a specialist other than the GP within both the public and private setting. However, if patients need to consult a specialist in secondary care, they often need a referral letter from a GP.

The challenge of the GP is to identify the few who have a potentially serious disease from among the many patients that consult for symptoms and signs signalling it (Figure 2). This applies especially to cancer. At the same time, a GP diagnoses only a handful of cancers annually, with each of colorectal, breast, lung and prostate. [23, 33, 34]

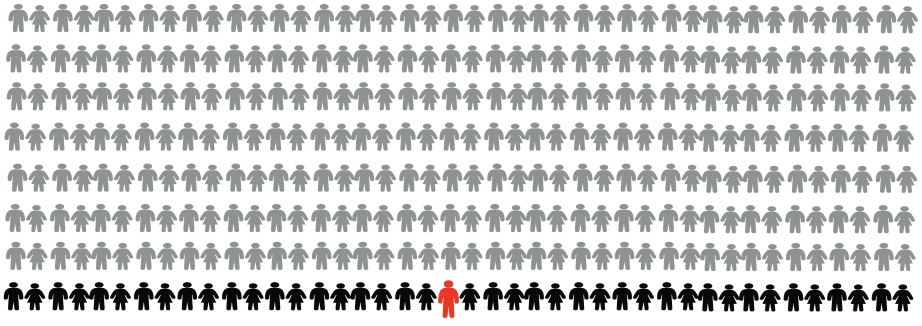

Figure 2. The number of patients that consult a GP before one single patient is diagnosed with cancer. Grey: All patients consulting a GP. Black: Patients with symptoms and signs of potential cancer Red: This patient is diagnosed with cancer. The figure is modified after original image from www.stockphoto.com

An increase in consultations with their GP is a pattern seen in cancer patients before diagnosis.[25, 35, 36] A large Danish study of cancer patients' use of primary care services showed that GP consultations doubled and diagnostic

investigations rose 10–11 times three months before cancer diagnosis.[37] There is a variation of the number of consultations in primary care depending on the type of cancer.[38-40] More consultations usually means a delay in cancer diagnosis. To better compare the measures of delay in primary care the term *primary care interval* is being used to defining the time between the first time a cancer patient presents with symptoms to a GP and their first referral for further investigation.[41-43]

Alarm symptoms versus non-specific symptoms

Alarm symptom are warning signs that demand action from the GP. Haemoptysis, for example, is considered a classical alarm symptom for lung cancer, and a lump in breast for breast cancer. These symptoms are well known to GPs and demand prompt investigation and referral (Figure 3). Approximately 50% of patients subsequently diagnosed with cancer, present to the GP with alarm or 'red flag' symptoms and the other half with non-specific or general symptoms. [44, 45]

Figure 3. Wake-up call cancer

In a study from Denmark, in nearly 6% of all consultations in general practice (one patient each day), the GP suspected cancer or another serious disease, but in the end, only 10% of these cases were diagnosed with cancer or another serious disease.[46] Studies of warning signs of cancer in general practice in Norway have shown that in 12% of the GP consultations, patients presented with alarm symptoms. In 24 % of these, the GPs suspected cancer, but less than 4% of these patients actually had cancer.[34, 47] If the initial symptoms of cancer were alarm symptoms, this might improve the outcome for the patients, because their recognition as potential cancer signs often results in investigation and referral by the GP, resulting in earlier cancer diagnosis.[19, 28, 45, 48]

Nevertheless, having alarm symptoms does not necessarily mean that the patients have a malignant disease because these symptoms are common in the general population. Another study indicated that 13%–15% of a Danish population had experienced alarm symptoms of the breast, lung, colorectum and urinary tract at least once within the last 12 months.[49]

The other half of cancer patients that do not consult a GP because of alarm symptoms present with non-specific or general symptoms.[45, 50] Non-specific symptoms such as tiredness or fatigue are an even greater challenge for a GP to interpret, because these clinical features cannot easily be associated with a specific organ or disease. General symptoms are also frequently presented in countless numbers of benign conditions and diseases. GPs are trained to quickly decide, based on the symptoms presented, whether the patient requires instant action, a moderate pace of investigation or watchful-waiting. The decisions leading to a work-up depend not only on the medical history, the results from diagnostic tests and physical exams, but also on the knowledge the GP has about the patient. These include age, gender, risk factors and finally the GP's own knowledge, experience and 'gut feeling'.[51-53]

Symptoms and diseases common in the general population are also common in general practice, thus being aware of epidemiology is important for GPs when they assess the clinical state of the patient and possible differential diagnoses. This is why most GPs apply one of the first clinical rules learnt from medical school: 'When you hear hoofbeats, think of horses not of zebras'.

Another way to describe cancer symptoms is by their 'symptom signature'. Cancers with a narrow symptom signature such as lump in breast for breast cancer or haematuria in bladder cancer are examples of symptom presentations which tend to be recognized early on. Cancers such as colorectal and lung cancer can present with a broad symptom signature consisting of multiple symptoms of which only a few are alarm symptoms that are strongly predictive of cancer.[54] That is the main reason why cancers with a broad symptom signature are more difficult to suspect and therefore to diagnose in a timely way.

The risks of different clinical features being an indication of cancer are often expressed in positive predictive values (PPV). A feature with a PPV of 3%

means that the person exhibiting this symptom, sign or disease has a 3% risk of having cancer. The PPV depends on the prevalence of the disease in the population, thus the rarer the disease, the lower the PPV. Alarm symptoms are especially indicative of cancer. A PPV of 5 % has previously been proposed as a threshold for alarm symptoms, but studies have found that they may have a PPV as low as 2% and as high as above 10%, both alone or in combination.[19, 33, 48]

Non-specific symptoms have lower PPVs than those considered alarm symptoms. However, there are no absolute rules for excluding cancer based on the characteristics of the symptoms, and even benign symptoms can be signs of cancer. Many cancer patients do not have high-risk symptoms but instead 'low-risk-but-not-no-risk' symptoms.[21, 33] Cancer can also be asymptomatic and would be detected by screening or by chance if the patient underwent medical investigation for another reason.

Another important aspect of cancer diagnostics is that our knowledge of cancer symptoms mainly stems from secondary care data. Cancer patients in hospitals do not necessarily exhibit the same clinical picture as patients consulting in primary care, because they are a selected population and the aspect of time is involved. The growth of the malignant tumours over time changes the clinical features both in number and in their characteristics. An example of this is haemoptysis in lung cancer, where a PPV as high as 35% has been reported from the secondary care setting, while lung cancer patients in primary care who present with this as a single symptom have a PPV of just over 2 %.[55, 56] When assessing cancer risks for patients presenting in primary care, data should be derived from the unselected population in primary care.

Decision support for cancer in primary care

Risk assessment tools

Because GPs encounter patients with such a diversity of medical conditions, a number of decision support tools have been designed to guide them to a feasible choice of treatment. GPs are acquainted with a number of different tools to support their decisions in care management, such as those to calculate a patient's risks for fatal complications in cardiovascular diseases or the risk for fracture in patients with osteoporosis. Such support tools also exist for

calculating cancer risks in primary care. Numerous risk tools are available which predict either current or future risk of cancer diagnosis.[57]

The risk assessment tool (RAT) for cancer in primary care, which has been developed and is being used in the UK, is an algorithm that can be used to calculate the absolute risk that a patient has an undiagnosed cancer based on certain risk factors and current symptoms. RATs are designed to support GPs in deciding which patients require further investigation or referral. These tools exist for common cancers such as lung cancer and colorectal cancer and for rarer cancers such as haematological malignancies.[33, 58, 59]

QCancer is another risk prediction algorithm developed in the UK to identify an individual's absolute risk of having a number of common cancers in the next two years. It is based on alarm symptoms, general symptoms and risk factors.[60, 61] Another predictive model has been designed in Israel for detecting patients at risk for colorectal cancer at an earlier stage by analysing complete blood counts, age and sex.[62]

Guidelines

Apart from risk assessment tools, organizations in some countries, for example, the UK's National Institute for Health and Care Excellence (NICE), have developed guidelines for suspected cancer. The guidelines include recommendations on the symptoms and signs that warrant investigation and referral for suspected cancer.[63] A 3% threshold for PPVs warranting urgent referrals is applied.

Urgent referrals for suspected cancer

In some European countries, the increasing awareness of poor cancer outcomes has resulted in national initiatives focused on early cancer detection and shorter waiting times for cancer treatment. In the UK, this initiative, called two-week wait referral, applies to patients with certain symptoms, signs and risk factors who will profit from an urgent admission for examination to confirm or exclude the suspicion of cancer. [64-66]

In Denmark, where cancer patients had a poorer five-year relative survival than many other countries in Western Europe, cancer was proclaimed to be an acute disease. In 2008, this resulted in the implementation of cancer

patient pathways (CPP), a strategy to reduce wait time for patients for whom there is a reasonable suspicion of having cancer.[67] This approach has been successful; wait times have shortened and collaboration between levels of care has improved. A recent study of the effect on survival has found higher relative survival and lower mortality rates among symptomatic cancer patients diagnosed through primary care after the implementation of CPP.[68] In international comparisons, cancer care in Sweden is characterized by high survival rates, but long wait times.[29, 69] In 2009, the Swedish government launched a national cancer strategy, and in 2015, inspired by the Danish system, introduced standardized care pathways (in Swedish *Standardiserade vårdförlopp*).[69] The objectives of this initiative were three: reducing wait time, increasing patient satisfaction with cancer care, and reducing regional inequalities (Figure 4). The same year Norway started a similar programme.

Figure 4. Each day counts. Campaign picture from the implementation of standardized care pathways in Sweden. www.cancercentrum.se

At present 28 standardized care pathways, have been implemented in the Swedish health care system for the most common and more rare cancers. The final three standardized care pathways are to be implemented in 2018. The start of this care process is defined by 'reasonable suspicion of cancer' in either primary or secondary care and identified by a set of indicators (symptoms or signs) and tests, which are different for each cancer. The symptoms and signs are for all the pathways but one derived from national clinical cancer care guidelines, which are based on data from secondary care.

Times are specified for all diagnostic procedures. The pathways are standardized up to the start of treatment for cancer. GPs use the criteria for referring to these pathways as guidelines for when to suspect cancer. A recent report from the Swedish National Board of Health and Welfare (Socialstyrelsen) shows that the proportions of cancers diagnosed in the care pathways is high and that wait times have been reduced for patients referred in some, but not all, of the cancer pathways.[70] Because these pathways were introduced recently, not enough time has passed to identify any improved survival.

This thesis is based on data from 2011, thus before the implementation of the standardized care pathways in Sweden.

Colorectal cancer

Colorectal cancer is the third most common cancer worldwide with more than 1.3 million cases reported annually.[71] In Europe it is the second most common cancer with more than 447,000 patients diagnosed each year.[11, 71] In Sweden, it is the fourth most common cancer, and more than 6500 patients are diagnosed annually.[14] Patients diagnosed with non-metastatic colorectal cancer have a good survival outcome, but the risk of dying from metastatic colorectal cancer is high.[72, 73]

Figures from the Swedish Colorectal Cancer Register for 2016 indicate that for patients with colon cancer who had undergone elective surgery, the 5-year relative survival for Stage I was 99%, Stage II, 94%, Stage III, 76% and Stage IV, 32%.[72] For rectal cancer, regardless of the mode of surgery (elective or non-elective) the 5-year relative survival for Stage I was 93%, Stage II, 86%, Stage III, 69% and stage IV, 17%.[73] Thus detection of colorectal cancer at an early stage improves survival considerably.

Alarm symptoms of colorectal cancer are generally considered to be rectal bleeding, change in bowel habit, weight loss and anaemia.[19, 74-76] In the UK, NICE published new guidelines for suspected cancer in 2015, and found evidence from 25–30 studies on single symptoms for colorectal cancer.[63] However, only nine of them reported on the cardinal symptom of rectal bleeding combined with other symptoms, and only two of these reported on other combinations of symptoms.[75, 77] Thus, there are only a few multi-

symptom studies on colorectal cancer. RAT and Qcancer, the two risk prediction tools used in the UK for detecting colorectal cancer, are not designed to capture less advanced disease. At present, there are no tools that GPs can use to identify early stage colorectal cancer patients.

Lung cancer

Lung cancer is the most common cancer worldwide and also one of the deadliest.[71] It is the fourth most common cancer in Europe with more than 410, 000 new cases annually.[71] In Sweden 4194 patients were diagnosed with lung cancer in 2015 and 3626 died from it.[13, 14] The high mortality rate is due to both diagnosis at a late stage and delay in treatment.[7, 8, 78, 79] The relative 5-year survival rate for lung cancer in Sweden is 18%.[30] This low survival rate is due to more than 50% of all Swedish lung cancer patients being diagnosed at Stage IV, with a relative 5-year survival rate of 2.6%. When lung cancer is diagnosed at Stage I, the relative 5-year survival rate is 63.8%.[30] This high proportion of cancer patients with metastasized lung cancer at diagnosis occurs in other countries. In the UK, half of the lung cancer patients diagnosed in 2014 with a known stage were diagnosed at Stage IV.[32]

The RAT for lung cancer and Qcancer in the UK are two decision support tools designed for primary care use. However, because half of the lung cancer patients are diagnosed at Stage IV, it is doubtful whether these tools are able to detect cancer at early stages.

Screening of target groups has been discussed as a method for early diagnosis of lung cancer. Low-dose computed tomography (LDCT) in defined populations of high-risk persons has shown high sensitivity and acceptable specificity.[80] Results from different cancer screening trials have shown that up to 70% of screen-detected, non-small lung cancers were found in Stage I compared to around 15% in routine care.[81] LDCT is currently being used as screening for lung cancer in the US.

3 AIM

The overall aims of this thesis were to explore how general practitioners could identify common cancers in patients in primary care, at an early stage, and to design a risk assessment tool.

The specific aims underlying this thesis were to do the following:

- Identify early diagnostic profiles, such as diagnostic codes and consultation patterns of patients with the most common cancers.

- Identify the consultation profiles including potential missed diagnostic opportunities and clinical features of cancer patients who frequently consult GPs.

- Identify clinical features of non-metastatic colorectal cancer and design a risk assessment tool for it.

- Identify clinical features of non-metastatic lung cancer.

- Compare the clinical features in GPs' referral letters for chest X-ray with clinical features expressed as diagnostic codes in the regional health care database.

4 PATIENTS AND METHODS

When designing a study, a research methodology that can be applied best to the research question or hypothesis should be used. Because the aim of this thesis is to use databases to identify the early clinical features of primary care patients with common cancers before their cancer diagnosis, an observational retrospective approach was considered the most suitable methodology.

The thesis is based on four different studies (see Table 1).

4.1 Design and setting

All four studies are total population-based case-control studies, as they are based on data for all incident cancers diagnosed in one year in a specific region. Data were collected from both national and regional healthcare databases in Region Västra Götaland (RVG), which is situated in the southwest of Sweden and has 1.6 million inhabitants (17 % of the Swedish population) (Figure 5). This region has both rural and urban areas and is representative of the whole of Sweden. The RVG has both public and private primary healthcare care units. In 2011, the year from which data were collected, there were 197 primary healthcare units in RVG, 113 public and 84 private.[82]

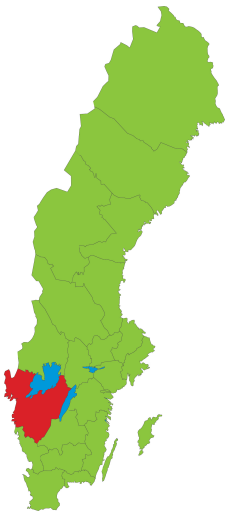

Figure 5. Sweden in green. Region Västra Götaland in red.

Table 1. Overview of the studies included in the thesis

Study/ Paper	I	II	III	IV
Design	Case-control	Case-control	Case-control	Case-control
Setting	Primary healthcare units in RVG	Primary healthcare units in RVG	Primary healthcare units in RVG	Primary healthcare units in RVG
Period	1 Jan 2010– 31 Dec 2011	1 Jan 2010– 31 Dec 2011	1 Jan 2010– 31 Dec 2011	1 Jan 2010– 31 Dec 2011
Study participants	4562 patients 17,979 controls	2570 patients 9424 controls	542 patients 2139 controls	373 patients 1472 controls
Data collection method	SCR, regional healthcare database	SCR, regional healthcare database	SCR, regional healthcare database	SCR, EIA, regional healthcare database
Primary outcome measures	Consultation frequency, symptom density by cancer type, OR for diagnostic codes	Consultation profiles and clinical features in patients with four or more GP consultations	PPV for clinical features, risk assessment tool non-metastatic colorectal cancer	OR for clinical features of non-metastatic lung cancer and clinical features in GPs' referral letters for chest X-ray

EIA= Enterprise Information Archive for radiology
OR = Odds ratio
PPV= Positive predictive value
RVG= Region Västra Götaland
SCR= Swedish Cancer Registry

4.2 Databases

The Swedish Cancer Registry

The Swedish Cancer Registry, (SCR) which was founded in 1958, is one of the oldest registries in Sweden and has high validity.[83] All physicians and pathologists in Sweden are obliged by law to report all incident cases of cancer in both living and dead patients to the registry.[14] 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

The regional healthcare database also called VEGA, is an administrative healthcare database which was established in 2000. It covers all hospitals, specialized outpatients 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.[84] Regular medical revisions have been made for this database for the diagnostic accuracy. At each consultation, physicians enter codes for patients' current disease or symptoms into the patients' medical records. 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 the regional healthcare database.

Enterprise information archive for radiology

In 2002, a decision was made to digitize all radiology departments in the region to better meet future needs. An enterprise information archive (EIA) was created for radiology information.[85] Both textual information and images can be shared in the region from the same virtual repository. This repository is one of the largest of its kind in the world.[86]

4.3 Diagnostic codes

Diagnostic coding is a tool that converts written information in medical records into codes that group and classify diseases, symptoms, disorders, pathological signs and abnormal findings. In Sweden, two main classification systems are used in primary health care. One is the Swedish version of the

International Classification of Diseases and Health problems 10th revision [ICD-10][87]. The other is the *Classification of Diseases and Health Problems 1997 Primary Care* [KSH97-P][88, 89], an abbreviated version of ICD-10, adapted to Swedish primary care to facilitate diagnostic coding. Physicians are obliged to enter codes for a patient's current disease(s) and symptoms into the patient's medical record at each consultation. Internationally, other classifications are used in primary care such as the International Classification of Primary Care (ICPC-2) for its better description of symptoms.[90]

The diagnostic codes used in all four studies, were registered when patients and their controls consulted their GP during the year preceding their cancer diagnosis. Because the controls had no cancer, their observation time corresponded to the observation time of their cases. We initially had more than 6000 different diagnostic codes and reduced their number according to clinical relevance. This was done by merging the ICD-10 four-character diagnostic codes to the closest three-character diagnostic codes. That resulted in 575 diagnostic codes (Figure 6).

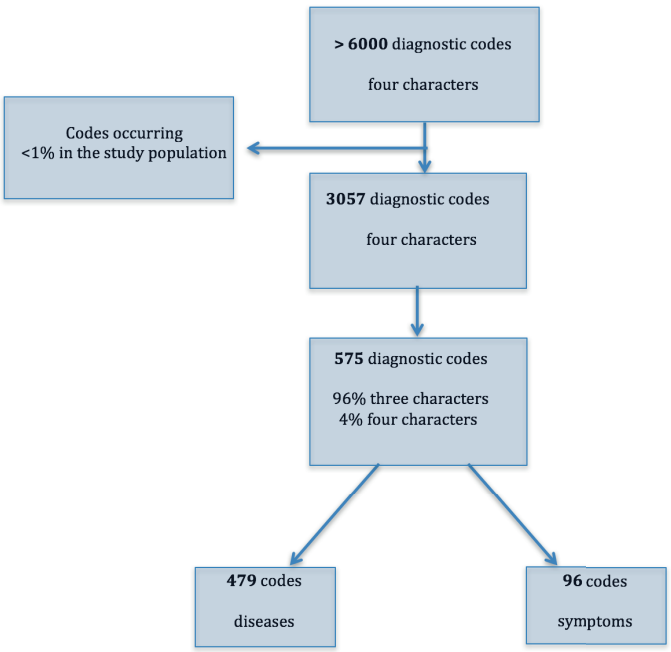

Figure 6. Flowchart of the merging process of diagnostic codes

The three-character codes are the core classification and the mandatory level for reporting to the World Health Organization's (WHO) mortality database[91] and for general international comparisons.

4.4 Data collection

Data were collected from the SCR and all cancer patients who were diagnosed in RVG in 2011 with the seven most common cancers: prostate, breast, colorectal, lung, gynaecological, and skin cancers including malignant melanoma were identified. These cancers constituted more than half of the annual cancer incidence in Sweden that year. The dates of the cancer diagnoses were retrieved for all cancers and for colorectal and lung cancer also stage information was retrieved.

The controls were selected from the regional healthcare database among all adult patients that had consulted a GP in RVG during 2010-2011. From this population four controls who were not diagnosed with cancer were matched on each cancer patient.

The diagnostic codes and dates of consultations in primary care for both cases and their controls were collected from the regional healthcare database in RVG from the period 2010-2011.

The third source of data was the EIA for radiology from which GPs' very first referral letters (in the year prior to the cancer diagnosis) for chest X-ray, containing detailed clinical information were retrieved from 2010-2011.

4.5 Ethical approval

The Regional Ethical Review Board in Gothenburg has approved all study protocols (252-12), amendment T 1004-12.

4.6 Study population and methods

Paper I

Paper I reports the results from patients diagnosed in 2011 in RVG with the seven most common cancers in Sweden: prostate, breast, colorectal, lung, gynaecological, and skin cancers including malignant melanoma. The cases were identified in the SCR. In total 4562 patients were included in the study, 50 % were female and the median age at diagnosis was 68 years (28–98). The sample recruitment process and inclusion and exclusion criteria are shown in Figure 7.

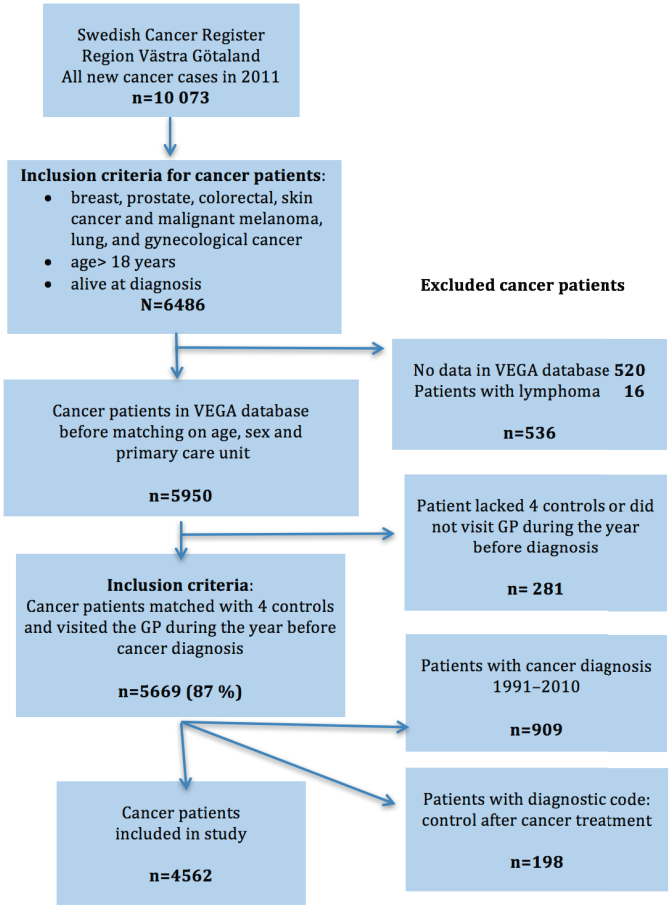

Figure 7. Flowchart of cancer patients included in the study

The controls were selected from VEGA. They had the same inclusion criteria as the cancer patients except for not being diagnosed with cancer. Four controls were matched to each case on age, sex and primary care unit. In total, 17,979 controls were included in the study.

The patients' unique personal identity numbers were linked to the VEGA and all the diagnostic codes and dates of consultations with a GP during 2010 and 2011 were retrieved for both cases and controls. The diagnostic codes for the consolidated diagnostic groups were used as variables for univariable conditional logistic regression. That resulted in a list of variables associated with each cancer type as well as their respective odds ratios (OR). We also calculated the lead time between consultation and cancer diagnosis and plotted consultation frequency expressed as weekly consultation frequency of cancer patients compared to controls. We also calculated symptom density expressed as weekly diagnostic code frequency. All analyses in both this and the other three studies were done in the statistical software R (version 3.0.1).

Paper II

Paper II also explored the diagnostic profiles of patients with the seven most common cancers, but in patients who had frequent consultations in primary care. They were diagnosed with their cancer in 2011 in the RVG and identified from the SCR. The inclusion and exclusion criteria were the same as in Paper I, except that only those that had consulted a GP four or more times in the year before their cancer diagnosis were included. Controls were identified in the regional healthcare database and had the same inclusion and exclusion criteria as the cases except for a cancer diagnosis. Four controls were matched to each patient, after which primary care data (including number of consultations) was obtained for those cases and controls. Whether a patient or control had four or more consultations was determined after the initial matching process, which means we simply retained all the patients and the controls who had four or more consultations. The median age of cases at diagnosis was 71(29-97) and median age of controls 70(29-97), 52 % of cases and 53% of controls were female. A total of 2570 cases and 9424 controls were finally included in the study.

The merged 575 diagnostic codes were then used for univariate conditional logistic regression at significance level 0.01. Those codes associated with

cancer were then analysed to see to which cancer they identified. The likelihood ratio (LR) was then calculated. LR is a measure that expresses the probability of any clinical finding in patients with a disorder divided by the probability of the same finding in patients without this disorder.[92]

The codes were then organized according to when in consultation order they were registered. Two groups were identified. One with early clinical features where some were registered at the two first consultations and less than 75% at the 4th or later GP consultation. The other group had more than 75% of the clinical features first presented at the 4th or later consultation. This was done to see if there might have been missed diagnostic opportunities at the first two consultations.

Paper III

Paper III identified all the patients in RVG that were diagnosed with colorectal cancer in 2011. Patients and matched controls were investigated for diagnostic profiles. Inclusion criteria were the same as in Paper I, including having a colorectal cancer with the stage registered. Exclusion criteria were also the same as in Paper I. However, in this paper, patients with metastasized colorectal cancer (Stage IV) were excluded as the aim was to study patients at Stages I–III (Figure 8).

A total of 542 patients with non-metastatic colorectal cancer were included in the study. The median age at cancer diagnosis was 72 years (30-94), 65% of the patients were female. Controls were generated from the regional healthcare database. Four controls were matched to each case on age sex and primary care unit, but 13 died before the diagnosis of their case. Included in the study were 2139 controls matched to patients with Stages I–III colorectal cancer. The unique personal identity numbers were linked to the regional healthcare database and all diagnostic codes and dates of consultations with a GP during 2010 and 2011 (one year before the date of cancer diagnosis) were retrieved for both cases and controls.

The 575 diagnostic codes 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

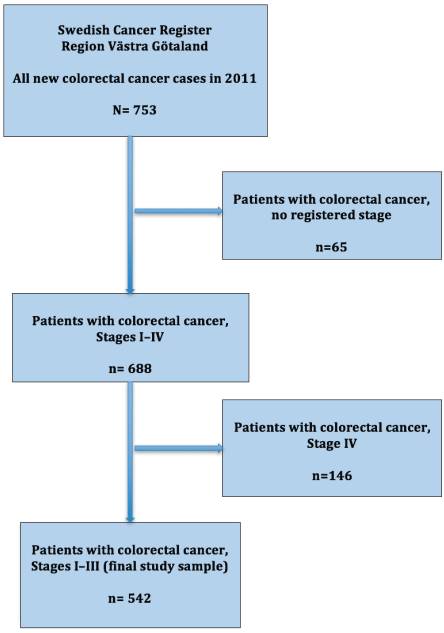

Figure 8. Sample recruitment flowchart

associated with cancer was compiled. A LR was then calculated for each variable (and combinations thereof). Using the LR, the incidence of colorectal cancer and Bayes' theorem[93], a PPV was calculated for each variable. This way we obtained PPVs for not only single but also for combined variables.

Paper IV

In Paper IV, all patients diagnosed in 2011 with lung cancer in the RVG were identified from the SCR. Because the study was based on the total population, no sample size was calculated. The aim was twofold: to identify the clinical features of non-metastatic lung cancer patients and to compare the clinical features in GPs' first referral letters for chest X-ray with clinical features expressed as diagnostic codes in the regional healthcare database. Therefore, two populations were studied.

The first population contained patients with non-metastatic lung cancer. The second was lung cancer patients that had been referred by a GP for a chest X-

ray and for whom an eligible GPs' referral letter was available from the EIA database or other repositories. The letter had to contain clinical information with symptoms and signs from physical examinations.

The inclusion and exclusion criteria for the first population were the same as for the cancer population in Study I, except that eligible patients had been diagnosed with lung cancer, and those with Stage IV lung cancer were excluded. In total, of the 373 patients with lung cancer that were identified in the SCR, 132 had Stages I–III (35%) non-metastatic cancer and the remaining 241 patients had Stage IV (65%).

Controls were selected from the regional healthcare database; the inclusion criteria were the same as for the patients with cancer, with the exception of a cancer diagnosis. Four controls had been matched to each case for age, sex and primary care unit but because 20 died before their cases received a cancer diagnosis, a total of 1472 controls were available. The unique personal identity numbers of both cases and controls were linked to the regional healthcare database, and data concerning diagnoses and dates of consultations with a GP during 2010 and 2011(one year before the date of the cancer diagnosis) were collected. The merged 575 diagnostic codes were used as variables for univariable conditional logistic regression. Variables found to be associated with cancer entered multivariable analyses, after which a list of statistically significant variables associated with lung cancer was compiled.

A review of the second population for which data was derived from the EIA showed that 151 out of 373 lung cancer patients had been referred by a GP for a first chest X-ray in the year prior to cancer diagnosis. The 151 GPs' referral letters for chest X-ray, which contained detailed clinical information with risk factors, symptoms and signs from physical examinations and pathological laboratory results, were retrieved either from the EIA database or other repositories. Two medical oncologists and a GP coded, independently of each other, the clinical features in all the referral letters for chest X-ray, using the ICPC-2 codes, because they are more symptom based. These codes were then compared with the ICD-10 diagnostic codes from medical records in the healthcare database.

5 RESULTS

5.1 Main results

- Both the frequency of GP consultations and number of diagnostic codes rose in tandem 50–100 days before a cancer diagnosis.

- More than half of the cancer patients consulted a GP four times or more before a cancer diagnosis. Features associated with cancer were presented early; they were focal and had benign characteristics.

- A certain combination of clinical features could be used to identify patients with non-metastatic colorectal cancer.

- Patients with non-metastatic lung cancer were not easily identified by clinical features.

- Clinical features in GPs' referral letters for chest X-ray were more frequent than corresponding features in the healthcare database.

5.2 Paper I

This paper studied early diagnostic profiles such as diagnostic codes and consultation patterns of cancer patients. Lump in breast, neoplasm of uncertain behaviour and abnormal serum enzyme levels were the diagnostic codes with highest OR. In cancers that presented with alarm symptoms such as palpable or visual changes, the numbers of consultations and diagnostic codes started to rise 50-60 days before cancer diagnosis, while cancer with less specific symptoms or signs such as those of the prostate, colorectum and lung had a rising trend of consultation frequency between 80–100 days (Figure 9).

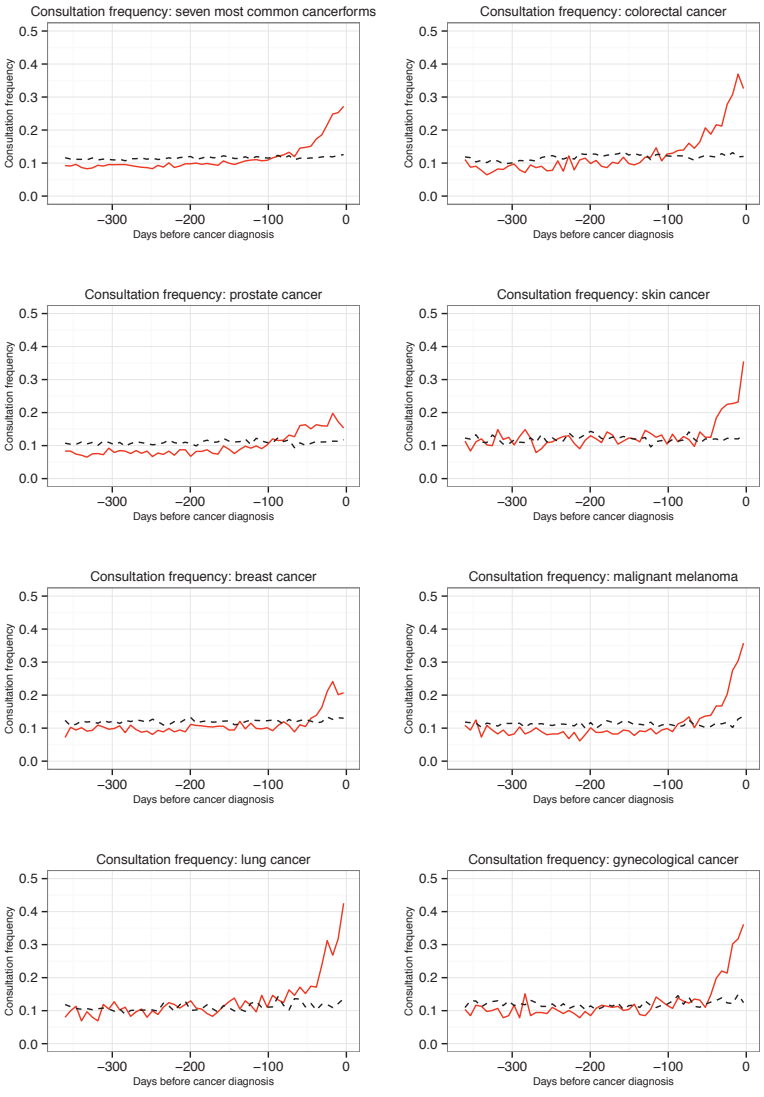

Figure 9. Consultation frequency: weekly consultation frequency of cancer patients (red continuous line) compared to controls (black interrupted line)

5.3 Paper II

Paper II looked for the consultation profile including potential missed diagnostic opportunities and clinical features of cancer patients with frequent consultations. It reports that 56% of patients with the seven most common cancers consulted a GP at least four times in the year before a cancer diagnosis. Among patients with breast cancer, the proportion was 48 %, colorectal cancer patients, 65 % and lung-and skin cancers, 66 %. The majority of clinical features associated with cancer were registered at the fourth or later consultation, and 60% with the highest LR were alarm symptoms. However, alarm symptoms formed only part of 40 % of the most prevalent codes. One out of six features associated with cancer or 17%, were presented at the two first consultations. These early clinical features were potential cancer signs, but not recognized as such. There were three kinds of features: alarm symptoms, for example, iron deficiency anaemia; potential cancer signs, such as abnormal serum enzymes and/or plasma protein levels and change in bowel habit; and focal benign disease from the prostate, digestive system or skin. These patients had to revisit a GP two more times or more often before being diagnosed with cancer.

5.4 Paper III

Paper III examined clinical features of non-metastatic colorectal cancer and described the design of a risk assessment tool. Five features were associated with non-metastatic colorectal cancer before diagnosis: bleeding, including rectal bleeding, melaena, and gastrointestinal bleeding PPV 3.9%(95% confidence interval [CI] 2.3–6.3); anaemia PPV 1.4%(95% CI 1.1–1.8); change in bowel habit PPV 1.1% (95% CI 0.9–1.5); abdominal pain PPV 0.9%(95% CI 0.7–1.1); and weight loss PPV 1.0%(95% CI 0.3–3.0); all P-value <0.05. The combination of bleeding and change in bowel habit had a PPV of 13.7% (95% CI 2.1–54.4); for bleeding combined with abdominal pain this was 12.2% (95% CI 1.8–51.2). A risk assessment tool for non-metastatic colorectal cancer was designed (Figure10).

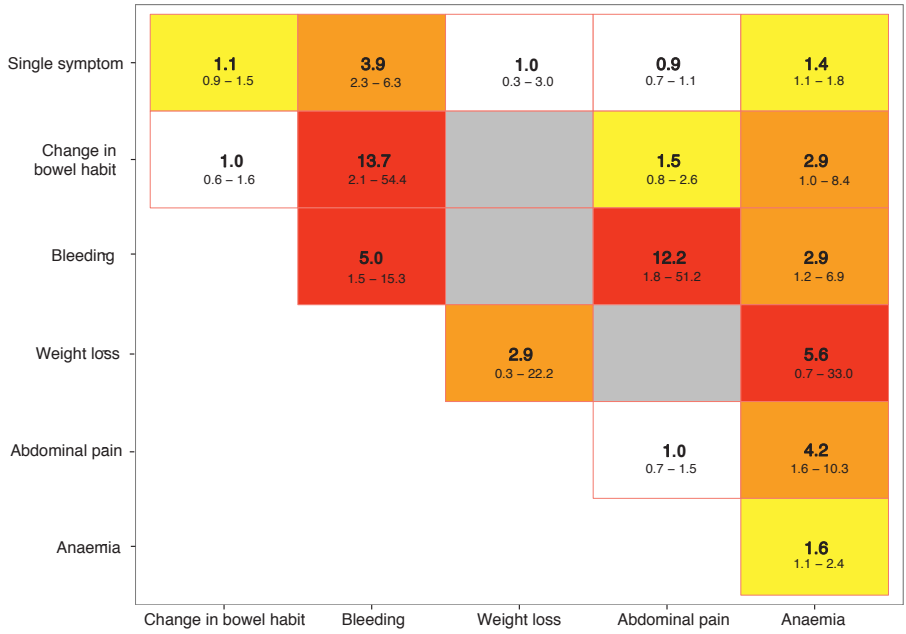

Figure 10. Risk assessment tool for non-metastatic colorectal cancer. Risk plot with PPV for colorectal cancer Stages I–III, in patients aged ≥50 years (against a background risk of 0.25%). Top-row single symptoms show the individual risk of each symptom. The diagonal rows show the PPV when the symptom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

5.5 Paper IV

This paper reports on the clinical features of non-metastatic lung cancer and comparison of data from GPs referral letter for chest X-ray and the regional healthcare database. The clinical features with the highest OR for non-metastatic lung cancer were vitamin B12 deficiency anaemia OR 6.7 (95% confidence interval [CI) 1.6–27.9), dyspnoea OR 5.0 (95% CI 2.0–12.7), and chronic bronchitis OR 5.0 (95% CI 1.3–18.6) (Table 2). Symptoms and diseases of the respiratory system were common in patients with both metastatic and non-metastatic lung cancer; however, the first group had more severe health conditions such as pulmonary embolism. Haemoptysis often seen as a risk marker for lung cancer was only seen in patients with metastatic disease.

Table 2. Univariable analysis of diagnostic codes in patients depending on lung cancer stage

M0=Stage I–III			M1=Stage IV		
ICD-10 code and diagnosis	**Prevalence %**	**OR(95%CI) ***	**ICD-10 code and diagnosis**	**Prevalence %**	**OR(95%CI) ***
D51 Vitamin B12 deficiency anemia	3.8	6.7 (1.6–27.9)	L20 Atopic dermatitis	1.2	12.0 (1.2–115.4)
R060 Dyspnoea	8.4	5.0 (2.0–12.7)	R042 Haemoptysis	2.1	9.6 (1.9–49.7)
J42 Unspecified Chronic bronchitis	3.8	5.0 (1.3–18.6)	I26 Pulmonary embolism	1.7	8.0 (1.5–43.7)
J44 Chronic obstructive pulmonary disease	20.6	4.3 (2.4–7.5)	W00 Fall due to ice and snow	1.7	8.0 (1.5–43.7)
I73 Other peripheral vascular diseases	4.6	4.2 (1.3–13.9)	M05 Rheumatoid arthritis with rheumatoid factor	2.9	4.7 (1.6–13.9)
B34 Viral infection of unspecified site	5.4	4.0 (1.4–11.4)	N20 Calculus of kidney and ureter	2.1	4.5 (1.2–17.1)
R05 Cough	13.8	3.8 (2.0–7.5)	W19 Unspecified fall	2.5	4.3 (1.3–14.3)
J18 Pneumonia	12.2	3.2 (1.6–6.2)	J18 Pneumonia	9.5	3.8 (2.1–7.0)

OR=odds ratio calculated between cases and controls. Diagnostic codes with OR > 3.
CI=confidence interval
* *P* value <0.05

In total, 40% of the patients referred for their first chest X-ray from primary care had non-metastatic lung cancer compared to 30% when being referred from secondary care. Clinical features that were GPs' reasons for referral were almost three times more frequent in referral letters than corresponding diagnostic codes in the regional healthcare database (Figure 11).

Figure 11. Word cloud with ICPC-2 codes [90] describing clinical features in GPs' referral letters to chest X-ray. Created by and used with the permission of Lars Lindsköld.

***R05**=Cough **A23**= Risk factor, not otherwise specified*
***A91**= Abnormal investigation **R02**=Shortness of breath/ dyspnoea*
***R95**=Chronic obstructive pulmonary disease **A04**=Weakness/tiredness **T08**=Weight loss*

6 DISCUSSION

6.1 General discussion of the results

It is important to use primary care data to provide evidence for primary care decisions, such as when to refer for investigation. Despite the existing urgent cancer pathways, the clinical skills of the GPs are as important as ever.

The age of the data presented in the studies for this thesis, might be considered an issue. However, there is no reason to believe that clinical features presented by cancer patients to primary care physicians today are any different. The results presented in this thesis show that patients in primary care who were found to have common cancers exhibit certain characteristics. Some of the traits that are a part of the pre-diagnostic pattern in cancer patients are frequent consultation for any reason, frequent consultations with benign focal symptoms and signs, and a certain combination of clinical features from the bowel. Hence, patients with common cancers do have early diagnostic profiles and consultation patterns.

Frequent consultations

Findings from the first study (Paper I) show an increase in consultations in primary care 50–100 days prior to cancer diagnoses of seven common cancers. Increased consultations before a cancer diagnosis has been reported for both common and more rare cancers.[35-37, 94] Thus, there is a positive correlation between frequent consultations and the time to cancer diagnosis.

Biswas et al. estimated a mean symptom lead time (the time from the presentation of symptoms caused by cancer in primary care and the diagnosis of cancer) of between four to six months, and medians of between two and three months for lung and colorectal cancer.[95] This is comparable with our findings, but we were not restricted to only studying symptoms previously known to be caused by cancer but all the diagnostic codes reported. However, the time frame presented in our study differs from the results in a national population-based study of all incident cancers in Denmark. It reported an increase in GP consultation five to six months before diagnosis.[37] This difference could be explained by the GPs in Sweden taking action sooner in

the diagnostic process when their patients presented with symptoms for which cancer could be a differential diagnosis. The International Cancer Benchmarking Partnership [96] have presented results that were favourable for the Swedish primary care. A significant correlation was demonstrated between the readiness of primary care practitioners to investigate symptoms indicative of cancer and cancer survival rates in different countries and jurisdictions.[97]

In our study, we saw a simultaneous increase in both the number of consultations and diagnostic codes. The time between the start of the increasing consultations frequency and final diagnosis was the shortest for breast and gynaecological cancer, which may not be surprising, as they probably presented with alarm symptoms. Colorectal and lung cancer had the longest times from onset of increasing consultations to diagnosis. These cancers are often diagnosed late, at a more advanced stage.[7, 9, 98]

The increased consultation rate is a clear and distinct diagnostic profile of patients in primary care with common cancers. However, as we only had access to the diagnostic codes from the regional healthcare database and not direct access to the medical records we could not conclude whether the increase was only due to patients' initiative to revisit or the result of the GP's work-up process. However, unless it is triggered by a previously known disease, an increased consultation rate in primary care should result in suspicion of cancer and result in an urgent referral or investigation by the GP to confirm or exclude cancer.

In the second study (Paper II) more than half of the cancer patients consulted a GP four or more times in the year before cancer diagnosis. This contrasts with a UK study in which 82 % of patients with 18 common and rarer cancers were referred to secondary care after the first or second consultation.[39] Our paper reported that 48% of breast cancer patients and 52% of malignant melanoma patients had four or more consultations in primary care before diagnosis, in contrast with another large UK study where only 7% of patients with breast cancer and 10 % of melanoma patients had three or more pre-referral consultations in primary care before being referred to a hospital to diagnose cancer.[38] Comparing different studies on pre-referral consultations in primary care before cancer diagnosis can be confusing, as

some relate to all consultations and others only to those with well-known cancer related symptoms. The fact that 94 % of all breast cancer patients in our study received the diagnostic code unspecified lump in breast at the fourth or later consultation, leads us to believe that the previous consultations were for other reasons than suspected cancer. We also believe that once the lump was registered, the patient was swiftly referred to secondary care for diagnosis.

In our study two out of three lung cancer patients consulted a GP four or more times in the year before their cancer diagnosis. A Danish study reported a similar pattern to ours regarding the consultation pattern in primary care in lung cancer patients. More than 72% of the patients had five or more consultations in the 12 months before the diagnosis.[36] Our findings that 65% of colorectal cancer patients consulted a GP four or more times in the year before cancer diagnosis is in line with another Danish study that found that almost 50 % of colorectal cancer patients consulted a GP five or more times during the year preceding colorectal cancer diagnosis.[35]Another study presenting results concordant with ours reported that when any reason for consultation was considered (as in this study), about three quarters of colorectal cancer patients had four or more pre-referral consultations in primary care.[99]

The findings presented in Paper II confirm what has been reported in literature, that half of the pre-diagnostic features are being presented to GPs as alarm symptoms. However, our findings differ in the sense that the patients that did not present with alarm symptoms did not present with diffuse symptomatology such as pain or fatigue, which has been reported elsewhere. [34, 44] Instead, they presented with focal features with benign characteristics. Thus, two more consultations were needed before cancer was diagnosed. A recently published study on diagnosing colorectal cancer in Dutch primary care concludes that GPs need to be more aware of repeated pre-existing complaints that are not alarm symptom, because these could account for missed diagnostic opportunities.[100] This is also in accordance with our findings. The clinical features that were presented early in patients with frequent consultations were attributed to abnormal blood tests, diseases of the digestive system, symptoms from the prostate and bladder and skin lesions.[101] These findings suggest that cancer symptoms and signs are presented early by

the patient, but are not recognized as such by the GP or that another consultation is needed for investigation and diagnostic tests. Thus, one out of six or 17% of these features might represent missed diagnostic opportunities.

Clinical features of non-metastatic colorectal cancer

To our knowledge the risk assessment tool for colorectal cancer presented in Paper III is the first tool for GPs to use to target patients with non-metastatic colorectal cancer.[102]

The RAT for colorectal cancer developed in the UK has PPVs for single symptoms and for pairs of symptoms that are similar to those outlined in our study. With the help of our tool, patients with increased risk of colorectal cancer could be identified at earlier stages if they have bleeding combined with diarrhoea, constipation or change in bowel habit. The PPV 13.7% is four times higher than this combination in the existing colorectal RAT in the UK.[75, 103] This is most probably due to the clinical features diarrhoea, constipation or change in bowel habit being merged into a single variable 'change in bowel habit' in out RAT. This was done as it was not possible to classify the diagnostic code 'change in bowel habit' as either diarrhoea or constipation and with the intention of making an easy to use RAT for GPs without needing to enter too many variables. The combination of bleeding and abdominal pain with PPV 12.2% in our RAT is also almost four times higher than the corresponding combination in the UK RAT. That difference can also be explained by the merging in our study of several diagnostic codes for the variable bleeding; it included colorectal bleeding, melaena, gastrointestinal bleeding and unclassified bleeding. The merging of codes could, of course, also be considered a limitation.

In a systematic review, the aim of which was to investigate the diagnostic value of symptoms for colorectal cancer in primary care, the summary estimated PPV of rectal bleeding and change in bowel habit was 11.8%.[104] This is similar to the findings in Paper III. However, that review did not study only non-metastatic colorectal cancer. Another systematic review from primary care settings which included only studies with non-metastatic colorectal cancer, found PPVs of 9%–12% with these combined symptoms.[48]

When comparing the PPVs of specific symptoms such as rectal bleeding, abdominal pain, weight loss, and anaemia alone in Qcancer, [105, 106] the results are very similar to ours in Paper III. However, the QCancer algorithm makes no distinction in the prediction of early or metastasised colorectal cancer.

A recently published Swedish thesis on diagnosing colorectal cancer in primary care states that the best test for detecting colorectal cancer is a combination of positive faecal immunochemical test (FIT) and/or anaemia.[107] We did not study the laboratory results of the participants in our study, which can be seen as a limitation. The most important issue addressed in Paper III is that our risk assessment tool for non-metastatic colorectal cancer in primary care will be able to identify patients with a potentially curable disease. A weakness of our RAT is that it has not yet been validated.

There are many potential benefits, as well as challenges, in the use of risk prediction tools for cancer in primary care and their implementations.[108] Further validation of different risk prediction tools is needed to assess the acceptability, clinical impact, and economic implications. [57]

Clinical features of non-metastatic lung cancer

The results from Paper IV show that non-metastatic lung cancer could not be identified by its clinical features.[109] The features with the highest OR were vitamin B12 deficiency anaemia, dyspnoea and chronic bronchitis. To our knowledge this is the first study to present clinical features of lung cancer patients with a non-metastatic disease. This is also the first study to present vitamin B12 deficiency anaemia as being a risk marker for non-metastatic lung cancer. Perhaps this finding is a paraneoplastic phenomenon though previously published studies have shown that individuals with vitamin B12 deficiency anaemia are at increased risk for other cancers such as gastric cancers and blood malignancies.[110, 111] The lack of laboratory results in our study to validate the diagnoses of vitamin B12 deficiency anaemia is a limitation. A recent systematic review from the UK suggests that patients in primary care with thrombocytosis have an increased risk of several cancers, among them lung cancer. Our study was unable to show this as we lacked data on blood test results.[112]

We found that symptoms and diseases from the respiratory system were common and associated with lung cancer in patients with Stages I–III as well as Stage IV. However, no specific traits could be used to differentiate patients with Stage I–III lung cancer from patients with metastasised lung cancer. Nevertheless, patients with Stage IV lung cancer had more severe diseases such as pulmonary embolism. They were also registered with diagnostic codes for falling, which could be interpreted as a marker of comorbidity and greater frailty.[113, 114] Compared to a UK study that found nine clinical features associated with lung cancer, we found only two in common with the non-metastatic lung cancer group: dyspnoea and cough.[33, 56]

A prospective study of patients referred to secondary care with suspicious symptoms reported that haemoptysis is the strongest symptom predictor of lung cancer but it occurred in only a fifth of the patients.[115] However, the study lacked staging information for the patients that exhibited this feature. In our study, haemoptysis, which has been regarded as a major sign of lung cancer in both textbooks of medicine, literature on cancer signs, and also in the lung RAT [56] for primary care in the UK, was only seen in patients with metastasized lung cancer. This is not surprising considering that at least 50% of patients with lung cancer are diagnosed at Stage IV.[31, 32] Thus, this finding emphasises that haemoptysis is a late sign of lung cancer and cannot be used for early detection.

Referral letters for chest X- ray

In Paper IV we reported that the clinical information in GPs' referral letters for chest X-ray was extensive, in contrast to what has been reported in the literature.[116] By coding detailed clinical information with risk factors, symptoms and signs from physical examinations and pathological laboratory results reported by the GP to the radiologist, we could compare the GPs' reasons for the request with the diagnostic codes from the regional database.

Clinical features that were GPs' reasons for requesting chest X-rays were almost three times more frequent in referral letters compared to the corresponding diagnostic codes in the regional healthcare database.

Even though all the patients in this study had consulted a GP in the year prior to their lung cancer diagnosis, there were differences in their diagnostic

profiles. These depended on whether they had been referred for their first chest X-ray by their GP or from secondary care. In total, 40% of the patients referred for their first chest X-ray from primary care had non-metastatic lung cancer compared to 30% when being referred from secondary care. This probably occurs because patients attending specialists in secondary care have more comorbidities, for example from the respiratory system. Patients suffering from chronic obstructive pulmonary disease (COPD) may be at risk of not being investigated in a timely way for lung cancer because their symptoms may be interpreted as being caused only by COPD.[36] An increasing number of studies emphasize that COPD is not only a risk factor for lung cancer but that the two diseases are closely linked by a number of factors, and that suffering from COPD increases the susceptibility for developing lung cancer.[117-119]

Risk assessment tools for cancer or screening?

Due to advances in both diagnostics and treatment of cancer the lives of many patients have been improved or saved. Nevertheless, the challenge is greater than ever to detect cancer at the earliest stage possible, so that patients can benefit from today's high medical standard in cancer treatment. Numerous countries have implemented screening of several common cancers. As mentioned, Sweden has screening for breast and cervical cancer and will likely soon implement colorectal cancer screening. The main advantage of screening is detection of cancer when it is still asymptomatic; at such an early stage the patient can obtain an efficient treatment with a higher likelihood of being cured.

We designed a risk assessment tool for non-metastatic colorectal cancer. However, we were not able to construct a similar tool for lung cancer. Because the clinical features of early and advanced lung cancer have very similar traits, differentiation of them by symptomatic presentation does not seem possible. Perhaps the thresholds for urgent investigation of suspected lung cancer has to be lower. A randomised controlled feasibility trial in the UK has studied patients at increased risk of lung cancer by using urgent chest X-ray referrals in patients presenting with new chest symptoms to primary care. Only 1.2% of expected 2.4% of trial participants were diagnosed with lung cancer.[120]

A number of screening studies on target risk population have reported a 70% detection rate of Stage I lung cancer compared with 15% in a non-screened population. So far, only one study has shown a reduced lung-cancer mortality with LDCT screening.[121] Lung cancer screening of target risk groups has been implemented in the US, but the results have been discouraging so far, because less than 4% of the eligible 6.8 million smokers in the US have received LDCT screening.[81, 122] Europe has yet not implemented lung cancer screening, and the final results of the NELSON trial, where the primary endpoint is a reduction in lung-cancer specific mortality is still awaited. However, a recently published European Union position statement urges the European countries to start planning for implementation of LDCT lung cancer screening as soon as possible.[123] Because of the lack of evidence on whether lung cancer can be detected at an early enough stage by its symptomatic presentation, today LDCT screening seems to be the only way of diagnosing patients with potentially curable lung cancer at an early stage.

6.2 Discussion in relation to methodology

Design

Case-control is the design of all four studies. This design was considered the best suited to the overall and specific aims. An alternative design could have been prospective cohort studies; but that was not feasible because of a limitation in time and resources. In all four studies, we had assigned four controls to each case, which increased the statistical power. We had also matched the controls on age, sex and primary care unit, which enabled us to reduce bias due to confounding. However, matching can also be a disadvantage as the effect of matching factor cannot be studied.

The main strength of the studies presented in this thesis is that they are total population based. All adult patients in a large region in Sweden that were diagnosed in 2011 with the seven most common cancers, which covered more than half of the annual cancer incidence, were included. No sample size was calculated for the power of these studies as they were all population based.

Databases

The SCR has high validity. On the other hand, the regional healthcare register has not undergone validation to any extent other than that of regular revisions by the regional healthcare administration. The data retrieved from this database is used for the reimbursement system to primary care providers.[84] Because all the diagnostic codes retrieved from this database were not externally validated other than through the healthcare administration, this can be seen as a weakness of our studies. The lack of data on blood test results in our studies is another weakness.

Diagnostic codes

A strength of the thesis is that we retrieved all diagnostic codes from GPs' face-to-face consultations in the specified time periods, not only codes previously known from the literature to be associated with cancer. Also, all the diagnostic codes for cases and controls were registered before the cancer diagnosis (i.e. registered prospectively) and automatically retrieved, thus avoiding selection bias.

However, the use of diagnostic codes could also be considered a limitation, because not all the symptoms for which a patient consulted their GP would be recorded as a diagnostic code in their medical record. Important information about symptoms in the free text of the medical record can be lost. As a result, we do not know if our inability to identify features that discriminate between metastatic and non-metastatic lung cancer patients resulted from not enough diagnostic codes being registered for lung cancer patients or because we had too few subjects to detect a difference. That methodological issue has been observed in other fields of research in primary care databases.[124] Since it is mandatory for the Swedish GPs to code, an extensive and reliable amount of data is available. In our study, some diagnostic codes were likely prevalent because they had been repeatedly registered in the medical record during previous consultations. When registering diagnostic codes, GPs are encouraged to register disease codes prior to symptom codes, as the disease burden of the patients is based on the codes. That is probably the main explanation as to why 83% of the merged final diagnostic codes that our studies are based upon are disease codes.

Because we obtained more than 6 000 different diagnostic codes, we had to reduce their number by a merging process which resulted in 575 codes. This merging process was done according to the clinical relevance. The method of the coding and the occasional cases where the final coding was uncertain was discussed within the research team. However, when coding the GPs' reasons for requesting a chest X-ray (Paper IV), all coding was done independently by three coders. Where the codes were not consistent between the three coders, a consensus was reached on the final coding.

Associations between symptoms and stage

Most of the literature on symptom presentation in cancer patients does not take the stage at diagnosis into account. In cancer, as in other diseases the severity of the clinical features increases with the progression of the disease. Expediting a cancer diagnosis is important, with no regard to the disease stage. However, survival depends most often on the stage at diagnosis. To increase survival rates for people with cancer, one of the most important factors to be considered is early identification of patients with a potentially curable disease, as they might benefit most from the cancer treatment.

That is why we studied the clinical features in two different populations: those with less advanced disease (Stages I–III) and patients with disseminated disease (Stage IV) in both colorectal and lung cancer patients.

Smoking status and observation time

Our ethical approval included only register data but not the medical records of the subjects. Because information on smoking status is in the free text of the medical records, we were not able to retrieve this variable, which is another limitation in our study. Most cancer symptoms occur 3-6 months before the cancer diagnosis, but a longer observation time than ours, the year before cancer diagnosis may be needed.[35, 75, 95]

Data analyses

We used statistical methods that we considered the most appropriate for the specific aims, which included producing descriptive statistics, and doing univariable and multivariable conditional logistic regression.

In study/paper I, the consolidated diagnostic groups were used as variables for univariable conditional logistic regression with the outcome being cancer Yes or No. This gave us a list of variables associated with each cancer type as well as their respective OR. Using consultation dates from the primary care data, we calculated the lead time between consultation and cancer diagnosis and plotted the consultation frequency over time as well as symptom density.

In study/paper II, after univariable conditional logistic regression LR were calculated for variables associated with the different cancers. Because only patients/controls that had consulted a GP at least four times (in the year prior to cancer diagnosis) were included, and this was done after matching controls to patient, there are not four controls for all patients. After this procedure, the codes were sorted in consultation order and organized into two groups: early clinical features where a considerable proportion had been registered at the two first consultations, and late clinical features.

For study/paper III, both univariable and multivariable logistic regression were used and resulted in PPVs for single and combination of variables, which made it possible to create a RAT for non-metastatic colorectal cancer. To do that we had to merge several of the diagnostic codes for similar clinical features into 10 clinical groups. Variable bleeding included colorectal bleeding, melaena, gastrointestinal bleeding and unclassified bleeding. As both bleeding from the upper and lower gastro intestinal tract were merged into one variable, this could be considered a limitation. Another important merging of codes into one variable 'change in bowel habit' combined diarrhoea, constipation and change in bowel habit. Because this resulted in fewer variables in the RAT, we managed to acquire statistical significant results and made the RAT easier to use. However, the confidence interval is wide for some of the combinations of clinical features.

Finally, in study/paper IV, OR were calculated for variables associated with non-metastatic lung cancer. Both univariable and multivariable conditional logistic regressions were performed, but because we had too few cases for each combination of features, no calculation of PPV could be done. A univariate conditional logistic regression was also performed for variables associated with lung cancer for two groups. All patients had consulted a GP in the year prior to their lung cancer diagnosis, but the first group was patients that had been referred for their first chest X-ray from primary care

and the other was patients referred from secondary care. In the second part of the study we compared the clinical features in GPs' referral letters for chest X-ray with the clinical features in the regional healthcare database.

7 CONCLUSION

The majority of patients that were later diagnosed with a common cancer had consulted a GP. Patients with the most common cancers displayed certain diagnostic profiles both in clinical features and in consultation patterns, which should raise the suspicion of a potential common malignancy.

The main conclusions of this thesis are as follows:

- Increased consultation frequency in primary care for any symptom or sign, unless it is caused by a previously known disease, is a risk marker for cancer and should result in a swift investigation or referral to confirm or exclude cancer.

- A considerable proportion of clinical features associated with cancer were presented in the two first consultations by cancer patients with four or more pre-referral consultations. These early clinical features that were focal and had benign characteristics might have been missed diagnostic opportunities.

- Colorectal bleeding combined with diarrhoea, constipation, a change in bowel habit or abdominal pain are the most powerful predictors for non-metastatic colorectal cancer and should result in prompt colorectal investigation. A risk assessment tool for non-metastatic colorectal cancer was possible to design.

- Based on data from our study, patients with non-metastatic lung cancer could not be identified by their symptoms and signs.

- The clinical features appeared more frequently in GPs' referral letters for chest X-ray than the corresponding features for lung cancer patients registered as diagnostic codes in the regional healthcare database.

8 FUTURE PERSPECTIVES

Due to earlier, timelier cancer diagnoses and improved results from cancer treatments, the survival of cancer patients has improved considerably in recent decades. However, we need to do better at applying new knowledge from cancer research in the primary care setting to every day practice.

The risk assessment tool for non-metastatic colorectal cancer, one of the main results of this thesis, can be used as a tool in primary care.

This thesis has demonstrated the difficulty in differentiating a lung cancer patient at an early stage from one with disseminated cancer. Our findings are in keeping with the reasons for the increasing interest in LDCT lung cancer screening. The question is probably no longer *if* lung cancer screening will be implemented in countries other than the US but *when*, even though this might not be the diagnostic solution for the majority of lung cancer patients.

Finally, research from primary care needs to be taken into account when national boards and health care authorities create guidelines and urgent referral pathways for cancer patients. Hopefully this thesis will contribute to a better understanding of the challenge GPs encounter each day and will increase knowledge of how to identify patients in primary care with potential cancer.

ACKNOWLEDGEMENT

This thesis would not have been possible to accomplish without the help and support from many people. In particular, I would like to thank:

Jörgen Månsson, my supervisor, for your support, encouragement and belief in me as a researcher when I was taking my first baby steps into the scientific field. I am very grateful for all your guidance and our fruitful discussions that helped me grow and develop critical thinking.

Peter Naredi, my co-supervisor, for all your support, fast feedback, enthusiasm and inspiring questions that challenged me to apply a larger perspective on my work.

Nils Conradi, former director of the Regional Cancer Centre West, for offering me the opportunity to do research and accomplish this thesis.

Lars Lindsköld, co-author, for all your hours, days and weeks that you invested in helping me with data collection from the EIA and other repositories for radiology. Your enthusiasm, generosity and coaching attitude inspired and helped me a lot.

Chenyang Zhang, co-author and statistician, without you there would be no statistical analyses done of the large number of variables collected, and no thesis.

Szilard Nemes, co-author and statistician, for all your initial work to set up a research database, and your helpfulness.

Erik Holmberg, statistician, for your extraction of data from the Swedish Cancer Register.

Kristina Narbro and **Mona-Lis Dalbrekt,** for your help in the extraction of data from VEGA.

Thomas Björk-Eriksson, director of Regional Cancer Centre West, for all your encouragement, support and assistance in coding.

Andreas Hallqvist, for your assistance in coding.

Ann-Charlotte Waldenström, for all your wise advice both in research and life, all help and support when being my mentor during my PhD studies.

Caroline Olsson, for your encouragement, interest in my work and support with Ann-Charlotte in academic writing.

Eva Deutsch, for your kindness and administrative help.

My research fellow students in 'Kliniska forskarskolan', for all good times and laughter while merging into the exciting world of research.

All my colleagues at Regional Cancer Centre West, for your positive attitude and interest in my work. Special thanks to **Malin Samuelsson** that had to put up with me and all Mac and Word related questions in the year when we shared the same room. Many thanks also to **Marcus Marin** and **Berndt-Inge Petersson** for IT support.

My colleagues 'Regional process owners', for your support, encouragement and interesting discussions.

My friends for being there both in the ups and downs of life, forest walks, cozy dinners and inspiring discussions.

My mother, for all your love, support and encouragement. You have always been my biggest supporter. Education and knowledge are some of the most precious values in life-you taught me.

My daughter, for your love and being who you are. This thesis is dedicated to you and your grandmother.

This work was supported by a grant from the Assar Gabrielsson Foundation and by financial support from Regional Cancer Centre West.

REFERENCES

1. Neal RD. Do diagnostic delays in cancer matter? *Br J Cancer*. 2009; **101 Suppl 2**:S9-S12.

2. Rupassara KS, Ponnusamy S, Withanage N and Milewski PJ. A paradox explained? Patients with delayed diagnosis of symptomatic colorectal cancer have good prognosis. *Colorectal Dis*. 2006; **8(5)**: 423-429.

3. Iversen LH, Antonsen S, Laurberg S and Lautrup MD. Therapeutic delay reduces survival of rectal cancer but not of colonic cancer. *Br J Surg*. 2009; **96(10)**: 1183-1189.

4. Tabar L, Vitak B, Chen TH, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology*. 2011; **260(3)**: 658-663.

5. Lindholm E, Brevinge H and Haglind E. Survival benefit in a randomized clinical trial of faecal occult blood screening for colorectal cancer. *Br J Surg*. 2008; **95(8)**: 1029-1036.

6. Eriksson H, Lyth J, Månsson-Brahme E, et al. Low level of education is associated with later stage at diagnosis and reduced survival in cutaneous malignant melanoma: A nationwide population-based study in Sweden. *Eur J Cancer*. 2013; **49(12)**: 2705-2716.

7. Torring ML, Frydenberg M, Hansen RP, Olesen F and Vedsted P. 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.

8. Neal RD, Tharmanathan P, France B, et al. Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes[quest] Systematic review. *Br J Cancer*. 2015; **112(s1)**: S92-S107.

9. Torring ML, Murchie P, Hamilton W, et al. Evidence of advanced stage colorectal cancer with longer diagnostic intervals: a pooled analysis of seven primary care cohorts comprising 11 720 patients in five countries. *Br J Cancer*. 2017; **117(6)**: 888-897.

10. Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015; **136(5)**: E359-E386.

11. 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.

12. Socialstyrelsen [The National Board of Health and Welfare]. Statistics on cancer incidence 2015. 2017 [Available from: http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/20468/2017-1-20.pdf. (Accessed 20 Dec 2017).

13. Socialstyrelsen [The National Board of Health and Welfare]. Statistisk om dödsorsaker 2015 [Statistics on causes of death 2015]: Socialstyrelsen; 2016 [Available from: http://www.socialstyrelsen.se/publikationer2016/2016-8-3. (Accessed 10 March 2017).

14. Socialstyrelsen [The National Board of Health and Welfare]. Statistikdatabas cancerregistret 2015[Statistics on cancer incidence 2015]: Socialstyrelsen; 2017 [Available from: http://www.socialstyrelsen.se/statistik/statistikdatabas/cancer. (Accessed 11 Dec 2017).

15. Socialstyrelsen[The National Board of Health and Welfare]. Cancerincidens i Sverige 2011[Cancerincidence in Sweden 2011] 2012 [Available from: http://www.socialstyrelsen.se/publikationer2012/2012-12-19. (Accessed 1 Jan 2015).

16. Socialstyrelsen [The National Board of Health and Welfare]. Nationella screeningprogram[National screening programmes] 2017 [Available from: http://www.socialstyrelsen.se/riktlinjer/nationellascreeningprogram. (Accessed 6 Dec 2017).

17. Peto J, Gilham C, Fletcher O and Matthews FE. The cervical cancer epidemic that screening has prevented in the UK. *Lancet*. 2004; **364(9430)**: 249-256.

18. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med*. 2000; **343(22)**: 1603-1607.

19. Jones R, Latinovic R, Charlton J and Gulliford MC. Alarm symptoms in early diagnosis of cancer in primary care: cohort study using General Practice Research Database. *BMJ*. 2007; **334(7602)**: 1040.

20. Hamilton W. Five misconceptions in cancer diagnosis. *Br J Gen Pract*. 2009; **59(563)**: 441-447.

21. Scheel BI and Holtedahl K. Symptoms, signs, and tests: The general practitioner's comprehensive approach towards a cancer diagnosis. *Scand J Prim Health Care*. 2015; **33(3)**: 170-177.

22. Hansen RP. Delay in the diagnosis of cancer. Research Unit and Department of General Practice, Faculty of Health Sciences.University of Aarhus, Denmark. PhD Thesis. Aarhus. 2008

23. Månsson J. The diagnostic process of cancer from the general practitioner's point of view. Göteborgs Universitet.PhD Thesis. Göteborg. 1999.https://gupea.ub.gu.se/handle/2077/13126

24. Demagny L, Holtedahl K, Bachimont J, Thorsen T, Letourmy A and Bungener M. General practitioners' role in cancer care: a French-Norwegian study. *BMC Res Notes*. 2009; **2(1)**: 200.

25. Ewing M, Naredi P, Nemes S, Zhang C and Månsson J. Increased consultation frequency in primary care, a risk marker for cancer: a case–control study. *Scand J Prim Health Care*. 2016; **34(2)**: 2015-2212.

26. Union for International Cancer Control. *TNM Classification of Malignant Tumours, 8th Edition*. Wiley-Blackwell, 2016, p.272.

27. The Swedish National Quality Register for breast cancer. Annual report 2016. 2017 [Available from: http://statistik.incanet.se/brostcancer/. (Accessed 5 Dec 2017).

28. Dregan A, Moller H, Charlton J and Gulliford MC. Are alarm symptoms predictive of cancer survival?: population-based cohort study. *Br J Gen Pract*. 2013; **63(617)**: e807-812.

29. 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. *The Lancet Oncology*. 2014; **15(1)**: 23-34.

30. Regionala Cancercentrum. Lungcancer-Årsrapport från Nationella lungcancerregistret (NLCR) 2015[Lung cancer- Annual Report from the National Lung Cancer Register 2015] 2016 [Available from: http://cancercentrum.se/globalassets/cancerdiagnoser/lunga-och-lungsack/kvalitetsregister/rapport/nlcr_rapport_tom2015_korr161219.pdf (Accessed 3 Jan 2017).

31. Regionala Cancercentrum. Nationell lungcancerrapport för diagnosår 2012-2016[National Lung Cancer report for year of diagnosis 2012-2016] 2017 [Available from: http://cancercentrum.se/globalassets/cancerdiagnoser/lunga-och-lungsack/kvalitetsregister/rapport/nlcr_rapport_tom2016_171120_publicera.pdf (Accessed 5 Dec 2017).

32. Cancer research UK. Lung cancer incidence statistics 2014 [Available from: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer/incidence - heading-Three (Accessed 6 Nov 2017).

33. Hamilton W. The CAPER studies: five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients. *Br J Cancer*. 2009; **101 Suppl 2**: S80-86.

34. Ingebrigtsen SG, Scheel BI, Hart B, Thorsen T and Holtedahl K. Frequency of 'warning signs of cancer' in Norwegian general practice, with prospective recording of subsequent cancer. *Fam Pract*. 2013; **30(2)**: 153-160.

35. Hansen PL, Hjertholm P and Vedsted P. Increased diagnostic activity in general practice during the year preceding colorectal cancer diagnosis. *Int J Cancer*. 2015; **137(3)**: 615-624.

36. Guldbrandt LM, Moller H, Jakobsen E and Vedsted P. General practice consultations, diagnostic investigations, and prescriptions in the year preceding a lung cancer diagnosis. *Cancer medicine*. 2017; **6(1)**: 79-88.

37. Christensen K, Fenger-Gron M, Flarup K and Vedsted P. Use of general practice, diagnostic investigations and hospital services before and after cancer diagnosis - a population-based nationwide registry study of 127,000 incident adult cancer patients. *BMC Health Serv Res*. 2012; **12(1)**: 224.

38. Lyratzopoulos G, Neal RD, Barbiere JM, Rubin GP and Abel GA. Variation in number of general practitioner consultations before hospital referral for cancer: findings from the 2010 National Cancer Patient Experience Survey in England. *Lancet Oncol*. 2012; **13(4)**: 353-365.

39. Lyratzopoulos G, Abel GA, McPhail S, Neal RD and Rubin GP. Measures of promptness of cancer diagnosis in primary care: secondary analysis of national audit data on patients with 18 common and rarer cancers. *Br J Cancer*. 2013; **108(3)**: 686-690.

40. Mendonca SC, Abel GA and Lyratzopoulos G. Pre-referral GP consultations in patients subsequently diagnosed with rarer cancers: a study of patient-reported data. *Br J Gen Pract*. 2016; **66(644)**: e171.

41. Allgar VL and Neal RD. Delays in the diagnosis of six cancers: analysis of data from the National Survey of NHS Patients: Cancer. *Br J Cancer*. 2005; **92(11)**: 1959-1970.

42. Hansen RP, Vedsted P, Sokolowski I, Sondergaard J and Olesen F. Time intervals from first symptom to treatment of cancer: a cohort study of 2,212 newly diagnosed cancer patients. *BMC Health Serv Res*. 2011; **11**: 284.

43. Weller D, Vedsted P, Rubin G, et al. The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. *Br J Cancer*. 2012; **106(7)**: 1262-1267.

44. Nielsen TN, Hansen RP and Vedsted P. [Symptom presentation in cancer patients in general practice]. *Ugeskr Laeger*. 2010; **172(41)**: 2827-2831.

45. Jensen H, Torring M, Olesen F, Overgaard J and Vedsted P. Cancer suspicion in general practice, urgent referral and time to diagnosis: a population-based GP survey and registry study. *BMC Cancer*. 2014; **14(1)**: 636.

46. Hjertholm P, Moth G, Ingeman ML and Vedsted P. Predictive values of GPs' suspicion of serious disease: a population-based follow-up study. *Br J Gen Pract*. 2014; **64(623)**: e346-353.

47. Scheel BI, Ingebrigtsen SG, Thorsen T and Holtedahl K. Cancer suspicion in general practice: the role of symptoms and patient characteristics, and their association with subsequent cancer. *Br J Gen Pract*. 2013; **63(614)**: e627-635.

48. Shapley M, Mansell G, Jordan JL and Jordan KP. Positive predictive values of >/=5% in primary care for cancer: systematic review. *Br J Gen Pract*. 2010; **60(578)**: e366-377.

49. Svendsen RP, Stovring H, Hansen BL, Kragstrup J, Sondergaard J and Jarbol DE. Prevalence of cancer alarm symptoms: a population-based cross-sectional study. *Scand J Prim Health Care*. 2010; **28(3)**: 132-137.

50. Vedsted P, Hansen RP and Bro F. [General practice and early cancer diagnosis]. *Ugeskr Laeger*. 2011; **173(24)**: 1712-1715.

51. Stolper E, Van Royen P, Van de Wiel M, et al. Consensus on gut feelings in general practice. *BMC Fam Pract*. 2009; **10**:66.

52. Johansen ML, Holtedahl KA and Rudebeck CE. How does the thought of cancer arise in a general practice consultation? Interviews with GPs. *Scand J Prim Health Care*. 2012; **30(3)**: 135-140.

53. Donker GA, Wiersma E, van der Hoek L and Heins M. Determinants of general practitioner's cancer-related gut feelings-a prospective cohort study. *BMJ open*. 2016; **6(9)**: e012511.

54. Koo MM, Hamilton W, Walter FM, Rubin GP and Lyratzopoulos G. Symptom Signatures and Diagnostic Timeliness in Cancer Patients: A Review of Current Evidence. *Neoplasia*. 2018; **20(2)**: 165-174.

55. Beckles MA, Spiro SG, Colice GL and Rudd RM. Initial Evaluation of the Patient With Lung Cancer[*]. *Chest*. **123(1)**: 97S-104S.

56. Hamilton W, Peters TJ, Round A and Sharp D. What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study. *Thorax*. 2005; **60(12)**: 1059-1065.

57. Usher-Smith J, Emery J, Hamilton W, Griffin SJ and Walter FM. Risk prediction tools for cancer in primary care. *Br J Cancer*. 2015; **113(12)**: 1645-1650.

58. Hamilton W, Peters TJ, Bankhead C and Sharp D. Risk of ovarian cancer in women with symptoms in primary care: population based case-control study. *BMJ*. 2009; **339**:b2998.

59. Cancer research UK. 2017 [Available from: http://www.cancerresearchuk.org/health-professional/diagnosis/suspected-cancer-referral-best-practice/risk-assessment-tool-rat?_ga=2.136796208.1672039831.1511961318-2038252476.1428484677. (Accessed 29 Nov 2017).

60. Hippisley-Cox J and Coupland C. Symptoms and risk factors to identify men with suspected cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract*. 2013; **63(606)**: e1-e10.

61. Hippisley-Cox J and Coupland C. Symptoms and risk factors to identify women with suspected cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract*. 2013; **63(606)**: e11-21.

62. 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.

63. National Institute for Health and Care Excellence. Suspected cancer: recognition and referral [NG12] 2015 [updated 2017 July. Available from: https://www.nice.org.uk/Guidance/NG12. (Accessed 30 Nov 2017).

64. Dodds W, Morgan M, Wolfe C and Raju KS. Implementing the 2-week wait rule for cancer referral in the UK: general practitioners' views and practices. *Eur J Cancer Care (Engl)*. 2004; **13(1)**: 82-87.

65. The National Health Service. Guide to NHS waiting time in England 2017 [Available from: https://www.nhs.uk/NHSEngland/appointment-booking/Pages/nhs-waiting-times.aspx. (Accessed 9 Dec 2017).

66. The National Health Service. The NHS Cancer plan 2000 [Available from: https://www.thh.nhs.uk/documents/_Departments/Cancer/NHSCancerPlan.pdf (Accessed 9 Dec 2017).

67. Olesen F, Hansen RP and Vedsted P. Delay in diagnosis: the experience in Denmark. *Br J Cancer*. 2009; **101 Suppl 2**:S5-8.

68. Jensen H, Torring ML and Vedsted P. Prognostic consequences of implementing cancer patient pathways in Denmark: a comparative

cohort study of symptomatic cancer patients in primary care. *BMC Cancer*. 2017; **17(1)**: 627.

69. Wilkens J, Thulesius H, Schmidt I and Carlsson C. The 2015 National Cancer Program in Sweden: Introducing standardized care pathways in a decentralized system. *Health Policy*. 2016; **120(12)**: 1378-1382.

70. Socialstyrelsen [National Board of Health and Welfare]. Standardiserade vårdförlopp i cancervården-lägesrapport 2017 [Standardized care pathways in cancer care- staus report], 2017 [Available from: http://www.socialstyrelsen.se/Lists/Artikelkatalog/Attachments/20764/2017-11-14.pdf (Accessed 30 Nov 2017).

71. World Health Organisation. International Agency for Research on Cancer. GLOBOCAN 2012: estimated cancer incidence, mortality and prevalence 2012 [Available from: http://globocan.iarc.fr/Pages/fact_sheets_population.aspx. (Accessed 11 Dec 2017).

72. Regionala Cancercentrum. Koloncancer-Nationell kvalitetsrapport för år 2016 från Svenska Kolorektalcancerregistret [National annual quality report 2016 from the Swedish Colorectal Cancer Register] 2017 [Available from: https://cancercentrum.se/globalassets/cancerdiagnoser/tjock--och-andtarm-anal/kvalitetsregister/rapporter-2017/kolon2016.pdf. (Accessed 10 Dec 2017).

73. Regionala Cancercentrum. Rektalcancer Nationell kvalitetsrapport för år 2016 från Svenska Kolorektacancerregistret [Rectal cancer national annual quality report 2016 from the Swedish Colorectal Cancer Register] 2017 [Available from: https://cancercentrum.se/globalassets/cancerdiagnoser/tjock--och-andtarm-anal/kvalitetsregister/rapporter-2017/rektal_kvalitetsregisterrapport_2016.pdf (Accessed 10 Dec 2017).

74. Fijten GH, Starmans R, Muris JW, Schouten HJ, Blijham GH and Knottnerus JA. Predictive value of signs and symptoms for colorectal cancer in patients with rectal bleeding in general practice. *Fam Pract*. 1995; **12(3)**: 279-286.

75. Hamilton W, Round A, Sharp D and Peters TJ. Clinical features of colorectal cancer before diagnosis: a population-based case-control study. *Br J Cancer*. 2005; **93(4)**: 399-405.

76. Hamilton W, Lancashire R, Sharp D, Peters TJ, Cheng K and Marshall T. The risk of colorectal cancer with symptoms at different ages and between the sexes: a case-control study. *BMC Med*. 2009; **7(1)**: 1-9.

77. Meineche-Schmidt V and Jorgensen T. 'Alarm symptoms' in patients with dyspepsia: a three-year prospective study from general practice. *Scand, J Gastroenterol*. 2002; **37(9)**: 999-1007.

78. Lyratzopoulos G, Wardle J and Rubin G. Rethinking diagnostic delay in cancer: how difficult is the diagnosis? *BMJ*. 2014; **349**: g7400.

79. Macleod U, Mitchell ED, Burgess C, Macdonald S and Ramirez AJ. Risk factors for delayed presentation and referral of symptomatic cancer: evidence for common cancers. *Br J Cancer*. 2009; **101 Suppl 2**(S92-S101).

80. Moyer VA. Screening for lung cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2014; **160(5)**: 330-338.

81. Heuvelmans MA, Groen HJ and Oudkerk M. Early lung cancer detection by low-dose CT screening: therapeutic implications. *Expert Rev Respir Med*. 2016: 1-12.

82. Regionkansliet Hälso och sjukvårdsavdelningen. Hälso-och sjukvård i Västra Götaland Verksamhetsanalys 2011[The regional health care analysis 2011 for Region Västra Götaland]. 2012.

83. Barlow L, Westergren K, Holmberg L and Talback M. The completeness of the Swedish Cancer Register: a sample survey for year 1998. *Acta Oncol*. 2009; **48(1)**: 27-33.

84. Bjorck S, Palaszewski B, Friberg L and Bergfeldt L. Atrial fibrillation, stroke risk, and warfarin therapy revisited: a population-based study. *Stroke*. 2013; **44(11)**: 3103-3108.

85. Lindsköld L. Designing and Using an Information Infrastructure in Radiology. Department of Clinical Science,Intervention and Technology.Karolinska institutet.PhD Thesis. Stockholm. 2012

86. Personal communication with Lindsköld L. Former director of the Enterprise Information Archive. 2017.

87. Socialstyrelsen[The National Board of Health and Welfare]. *Internationell statistisk klassifikation av sjukdomar och relaterade hälsoproblem-systematisk förteckning (ICD-10-SE)[Swedish version of the International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10)]*. Stockholm: Socialstyrelsen, 2010.

88. Socialstyrelsen [The National Board of Health and Welfare]. *Klassifikation av sjukdomar och hälsoproblem 1997. Primärvård (KSH97-P) [Classification of Diseases and Health Problems 1997. Primary Care (KSH97-P)]*. Stockholm: Socialstyrelsen, 1997.

89. Nystrom M, Vikstrom A, Nilsson GH, Ahlfeldt H and Orman H. Enriching a primary health care version of ICD-10 using SNOMED CT mapping. *J Biomed Semantics*. 2010; **1(1)**: 7.

90. Wonca International Classification Committee. International Classification of Primary Care-2nd edition 2003 [Available from: http://www.kith.no/upload/2705/ICPC-2-English.pdf (Accessed 27 April 2017).

91. World Health Organization. WHO mortality database 2016 [Available from: http://apps.who.int/healthinfo/statistics/mortality/whodpms/. (Accessed 12 Nov 2017).

92. McGee S. Simplifying Likelihood Ratios. *J Gen Intern Med*. 2002; **17(8)**: 647-650.

93. Papoulis A. *Bayes' theorem in statistics and Bayes' theorem in statistics (reexamined)*. 2nd ed. ed. New York, NY: McGraw-Hill, 1984, p.38-114.

94. Dommett RM, Redaniel MT, Stevens MCG, Hamilton W and Martin RM. Features of childhood cancer in primary care: a population-based nested case–control study. *Br J Cancer*. 2012; **106(5)**: 982-987.

95. Biswas M, Ades AE and 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.

96. Nicholson BD, Mant D, Neal RD, et al. International variation in adherence to referral guidelines for suspected cancer: a secondary analysis of survey data. *Br J Gen Pract*. 2016; **66(643)**: e106-e113.

97. Rose PW, Rubin G, Perera-Salazar R, et al. Explaining variation in cancer survival between 11 jurisdictions in the International Cancer Benchmarking Partnership: a primary care vignette survey. *BMJ open*. 2015; **5(5)**.

98. 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.

99. Renzi C, Lyratzopoulos G, Card T, Chu TP, Macleod U and Rachet B. Do colorectal cancer patients diagnosed as an emergency differ from non-emergency patients in their consultation patterns and symptoms? A longitudinal data-linkage study in England. *Br J Cancer*. 2016; **115(7)**: 866-875.

100. Brandenbarg D, Groenhof F, Siewers IM, van der Voort A, Walter FM and Berendsen AJ. Possible missed opportunities for diagnosing colorectal cancer in Dutch primary care: a multimethods approach. *Br J Gen Pract*. 2018; **68(666)**: e54.

101. Ewing M, Naredi P, Zhang C and Månsson J. Diagnostic profile characteristics of cancer patients with frequent consultations in primary care before diagnosis: a case-control study. *Accepted for publication in Family Practice Feb 2018*.

102. Ewing M, Naredi P, Zhang C and Månsson J. Identification of patients with non-metastatic colorectal cancer in primary care: a case-control study. *Br J Gen Pract*. 2016; **66(653)**: e880-e886.

103. Hamilton W, Green T, Martins T, Elliott K, Rubin G and Macleod U. Evaluation of risk assessment tools for suspected cancer in general practice: a cohort study. *Br J Gen Pract*. 2013; **63(606)**: e30-36.

104. Astin M, Griffin T, Neal RD, Rose P and Hamilton W. The diagnostic value of symptoms for colorectal cancer in primary care: a systematic review. *Br J Gen Pract*. 2011; **61(586)**: e231-243.

105. Hippisley-Cox J and Coupland C. Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract*. 2012; **62(594)**: e29-e37.

106. Hippisley-Cox J and Coupland C. Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. *BMJ open*. 2015; **5(3)**: e007825.

107. Högberg C. Diagnosing colorectal cancer in primary care. Department of Public Health and Clinical Medicine.Umeå university.PhD Thesis. Umeå. 2017.http://umu.diva-portal.org/

108. Walker JG, Licqurish S, Chiang PP, Pirotta M and Emery JD. Cancer risk assessment tools in primary care: a systematic review of randomized controlled trials. *Ann Fam Med*. 2015; **13(5)**: 480-489.

109. Ewing M, Naredi P, Zhang C, Lindsköld L and Månsson J. Clinical features of patients with non-metastatic lung cancer in primary care: a case-control study. *BJGP Open*. 2018.

110. Murphy G, Dawsey SM, Engels EA, et al. Cancer Risk After Pernicious Anemia in the US Elderly Population. *Clin Gastroenterol Hepatol*. 2015; **13(13)**: 2282-2289 e4.

111. Anderson LA, Gadalla S, Morton LM, et al. Population-based study of autoimmune conditions and the risk of specific lymphoid malignancies. *Int J Cancer*. 2009; **125(2)**: 398-405.

112. Bailey SER, Ukoumunne OC, Shephard E and Hamilton W. How useful is thrombocytosis in predicting an underlying cancer in primary care? a systematic review. *Fam Pract*. 2017; **34(1)**: 4-10.

113. Stenhagen M, Nordell E and Elmstahl S. Falls in elderly people: a multifactorial analysis of risk markers using data from the Swedish general population study 'Good ageing in Skane'. *Aging Clin Exp Res*. 2013; **25(1)**: 59-67.

114. Bandeen-Roche K, Seplaki CL, Huang J, et al. Frailty in Older Adults: A Nationally Representative Profile in the United States. *J Gerontol A Biol Sci Med Sci*. 2015; **70(11)**: 1427-1434.

115. Walter FM, Rubin G, Bankhead C, et al. Symptoms and other factors associated with time to diagnosis and stage of lung cancer: a prospective cohort study. *Br J Cancer*. 2015; **112 Suppl 1**: S6-13.

116. Triantopoulou C, Tsalafoutas I, Maniatis P, et al. Analysis of radiological examination request forms in conjunction with justification of X-ray exposures. *Eur J Radiol*. 2005; **53(2)**: 306-311.

117. Biswas A, Mehta HJ and Folch EE. Chronic obstructive pulmonary disease and lung cancer: inter-relationships. *Curr Opin Pulm Med*. 2017.

118. Sekine Y, Hata A, Koh E and Hiroshima K. Lung carcinogenesis from chronic obstructive pulmonary disease: characteristics of lung cancer from COPD and contribution of signal transducers and lung stem cells in the inflammatory microenvironment. *Gen Thorac Cardiovasc Surg*. 2014; **62(7)**: 415-421.

119. Vermaelen K and Brusselle G. Exposing a deadly alliance: novel insights into the biological links between COPD and lung cancer. *Pulm Pharmacol Ther*. 2013; **26(5)**: 544-554.

120. Neal RD, Barham A, Bongard E, et al. Immediate chest X-ray for patients at risk of lung cancer presenting in primary care: randomised controlled feasibility trial. *Br J Cancer*. 2017; **116(3)**: 293-302.

121. Aberle DR, Adams AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med*. 2011; **365(5)**: 395-409.

122. Jemal A and Fedewa SA. Lung cancer screening with low-dose computed tomography in the united states—2010 to 2015. *JAMA Oncology*. 2017.

123. Oudkerk M, Devaraj A, Vliegenthart R, et al. European position statement on lung cancer screening. *The Lancet Oncology*. 2017; **18(12)**: e754-e766.

124. 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(1)**: 105.