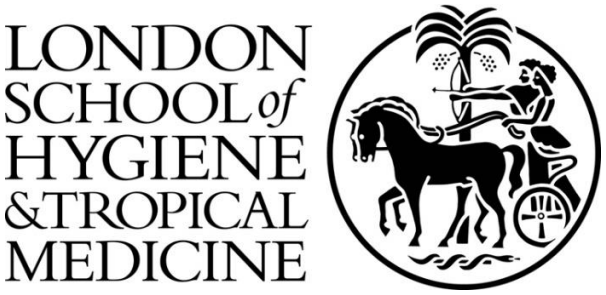


Statistical approaches for monitoring early cancer diagnosis in England

Patrick Muller

Thesis submitted in accordance with the requirements for the degree of Doctor of Philosophy of the University of London

April 2021

Department of Non-Communicable Disease Epidemiology
Faculty of Epidemiology and Population Health
London School of Hygiene & Tropical Medicine

No funding received

Research group affiliations:
Inequalities in Cancer Outcomes Network,
Cancer Survival Group

Supervisors

Dr Laura Woods

Department of Non-Communicable Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene & Tropical Medicine

Dr Sarah Walters

Department of Population Health

Faculty of Epidemiology and Population Health

London School of Hygiene & Tropical Medicine

Professor Michel Coleman

Department of Non-Communicable Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene & Tropical Medicine

Advisory Team

Dr Aurelien Belot

Department of Non-Communicable Disease Epidemiology

Faculty of Epidemiology and Population Health

London School of Hygiene & Tropical Medicine

Professor James Carpenter

Department of Medical Statistics

Faculty of Epidemiology and Population Health

London School of Hygiene & Tropical Medicine

Declaration of Authorship

I, Patrick Muller, confirm that the work presented in this thesis titled 'Statistical approaches for monitoring early cancer diagnosis in England' is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: P MULLER

Date: 10/05/2021

Acknowledgements

Projects conducted in the Cancer Policy Programme of the LSHTM Cancer Survival Group motivated much of the work that would later be done in this thesis. I'm grateful to all members of that programme for providing a stimulating and supportive environment: Sara Benitez Majano, Melanie Morris, Chiara Di Girolamo, Bernard Rachet, and Michel Coleman. Many other colleagues at LSHTM also generously provided support and advice: Aurelien Belot advised on methodology throughout the PhD; James Carpenter provided feedback on the upgrading report and missing data handling; Matteo Quartagno advised on the implementation of *jomo* in R; and Edmund Njeru Njagi advised on multilevel modelling.

Others helped the project indirectly: Caroline Doré, Nick Freemantle, and Jan van der Meulen have all supported my career these past few years, I'm grateful to them for helping me to balance the PhD with other commitments.

Laura Woods helped keep the project on track through some difficult times, and I'm indebted to her for her constant support, insight, pragmatism, and patience. Sarah Walters also saw it evolve from the start to the end, first as an inspiring manager, then later as a conscientious and supportive supervisor, and I'm extremely thankful to her for making it possible. I couldn't have asked for better supervisors!

I'd finally like to thank my family. My brother Matthew and sister Louise saw me through plenty of difficult times these past 30-odd years. Chris joined me on this PhD journey and made it her own too, and I'm eternally grateful for her love, support, and encouragement. Finally, I owe thanks to my mother Michèle, who taught me the importance of tenacity, kindness, and good humour in the face of adversity.

Abstract

Background

Increasing early-stage diagnosis is a priority of health policy in England. Numerous interventions to effect improvements were implemented during 2008-2013, including symptom awareness campaigns and bowel cancer screening. There were also structural changes to health services, with smaller increases in health spending from 2010 and a reorganisation in 2013.

An analysis of early diagnosis trends and geographic inequalities is needed to assess the impact of these changes. There are challenges to monitoring early diagnosis trends during 2008-2013, and afterwards, however. Disease stage was not recorded in the cancer registrations of many patients. These patients have poorer outcomes, and assuming early diagnosis was as common for them as patients whose stage was recorded may introduce bias. Case-mix factors may bias comparisons of health services performance, and comparisons between local areas can be limited by sparse data. Finally, other indicators alongside stage can be used in monitoring. Consideration of the added value and interpretation of these is merited.

These challenges have not previously been addressed in a national analysis of early diagnosis trends. The aim of this thesis is to address them, then apply the findings to evaluate trends and geographic inequalities during 2008-2013 for colorectal cancer, non-small cell lung cancer, and ovarian cancer. The implications for monitoring of early diagnosis after 2013 are then assessed.

Methods

Different early diagnosis indicators were described, and reasons for missing data were considered, with reference to the literature and data analysis. A conceptual framework for determinants of early diagnosis was created, and used to identify case-mix factors to be adjusted for in the substantive analysis. The association between different indicators and survival was evaluated through data analysis and a systematic literature review. Methods for case-mix adjustment, geographic comparisons, and handling missing data were surveyed. Finally, an analysis of early diagnosis trends and geographic inequalities during 2008-2013 was conducted using multilevel logistic regression, with multiple imputation to reduce bias from missing stage data. Sensitivity and simulation analyses were performed to assess the treatment of missing data.

Results

Stage was typically unreported for administrative reasons, but occasionally because the patient was frail; or died; or was treated privately. Improvements in staging data collection resulted in stage being less commonly missing for administrative reasons in 2013 than in 2008. Age, sex, comorbidities, and tumour morphology and topography were identified as key case-mix factors. Stage had clearest and most consistent interpretation of the indicators assessed. Multiple imputation was identified as the optimal approach to reduce bias from missing stage data.

There was evidence for an increase in the percentage of patients diagnosed at stages I or II for each of the three cancers analysed, including a step-change improvement for colorectal cancer (from 32% in 2008-09 to 44% in 2012-13). Geographic inequalities reduced. For ovarian cancer, estimated trends were different between analyses which did and didn't use multiple imputation. Sensitivity analyses indicated that the multiple imputation model was specified correctly, and that results were robust to some residual bias. It was found to be necessary to use information on one year's survival time to impute stage accurately.

Interpretation

Completeness of stage recording improved during 2008-2013, and afterwards to over 85% by 2018. Analysts may choose to disregard data on patients whose stage was not recorded, as they compromise a small proportion of the total. However, these patients have poorer outcomes, and the disparity between their outcomes and outcomes of patients with stage recorded has increased as stage recording has improved. Therefore, material bias may still be introduced from excluding them in analyses after 2013. Future evaluations, including of the overall impact of COVID-19 on early diagnosis, should use multiple imputation to account for patients whose stage is not recorded. This will give a more accurate picture of trends in early diagnosis of cancer in the whole population.

Contents

List of Tables

List of Figures

List of abbreviations and acronyms

2WW	Two Week Wait
A&E	Accident & Emergency
CCG	Clinical Commissioning Group
DH	Department of Health
FIT	Faecal immunochemical test
FOBT	Faecal occult blood test
GLM	Generalized linear model
HES	Hospital Episode Statistics
ICD	International Classification of Diseases
IMD	Indices of Multiple Deprivation
MAR	Missing at random
MCAR	Missing completely at random
MDT	Multidisciplinary team
MI	Multiple imputation
MNAR	Missing not at random
NCRAS	National Cancer Registration and Analysis Service
NICE	National Institute for Health and Care Excellence
NOS	Not otherwise specified
NSCLC	Non-small cell lung cancer
ONS	Office for National Statistics
PCT	Primary Care Trust
PHE	Public Health England
RCT	Randomised controlled trial
TNM	TNM Classification of Malignant Tumour

1. Thesis overview

1.1 Motivation

Prognosis of cancer is improved if it is diagnosed at an early stage [1]. Increasing the percentage of patients diagnosed at early stages has been a priority of recent cancer policy in England [2, 3]. Numerous interventions to effect improvements were implemented during 2008-2013, a period also characterised by a restructure of the National Health Service and increasing demand for services [4]. Data on early diagnosis in the whole population is needed to monitor the full impact of these changes. Robust evaluations are challenging, however. The indicators available in population-based data sources are not collected for the purposes of this monitoring and thus have complex interpretations. Data on some indicators is missing for many patients [5]. Handling of the missing data requires assumptions, which may be wrong. Case-mix factors need to be considered, and assessments of geographic differences may be compromised by sparse data. I aim to address these methodological challenges and apply the results in an analysis of stage trends during 2008-2013 for three cancers: colorectal cancer, non-small cell lung cancer, and ovarian cancer. Through this research I also aim to provide recommendations for accurate monitoring of early diagnosis monitoring after 2013.

1.2 Aims

1. To describe and interpret different early cancer diagnosis indicators in population-based data sources, and identify the most informative for monitoring early diagnosis trends;
2. To identify appropriate statistical methods to measure early diagnosis trends with minimal bias;
3. To evaluate changes in early diagnosis, and changes in geographic inequalities during the period 2008-2013 for three cancers (the “substantive analysis”);
4. To assess the potential for bias from missing data in the substantive analysis;
5. To provide recommendations for future monitoring of early diagnosis.

1.3 Objectives

1. With reference to the literature and data analysis, to explain how different early diagnosis indicators are derived and their meaning, including whether there is missing data for them in population-based data sources and if so, why;
2. To construct a conceptual framework for determinants of early diagnosis, and use it to identify confounding case-mix factors which should be adjusted for in analyses;
3. To identify an early diagnosis indicator for the substantive analysis which is associated with patient survival and which has a clear clinical interpretation, through a literature review and data analysis of patient records;

4. To review statistical techniques for case-mix adjustment and for evaluating geographic inequalities, identifying those which should be used in the substantive analysis;
5. To identify appropriate techniques to minimise bias from missing data;
6. Using the optimal indicator and statistical methods approaches previously identified, to produce statistics which measure changes in geographical inequalities, in early diagnosis in England during the period 2008−2013;
7. To consider whether the missing data multiple imputation model may be mis-specified, and if so how great the error would have to be to change analysis conclusions, and;
8. Considering the strengths and weaknesses of the methods surveyed above, results from the analysis, and recent developments in public health, to recommend how monitoring of early diagnosis should be conducted in future.

1.4 Thesis Structure

The thesis is organised into three introductory Chapters, followed by three research reports interspersed with Chapters on data items and statistical methods, leading to the overall discussion.

Changes to health services in England during 1990-2019 are described in Chapter 2. The impact of these, and of successive cancer plans and other interventions, on early diagnosis is considered. Chapter 3 introduces targets and performance measures as tools to ensure health services are performing, and discusses their risks and benefits. The added value of the work planned in this thesis is then explained in the context of the current landscape of cancer intelligence. The clinical background for diagnosis and treatment of the three cancers is discussed in Chapter 4.

In Chapter 5, the datasets I use to evaluate early diagnosis changes in England during 2008-2013 are introduced. The derivation of different patient- and tumour-level variables in these is explained. The interpretation of each variable is then considered, including the reasons for missing data on early diagnosis indicators (**Objective 1**).

A conceptual framework for the determinants of early diagnosis is presented in Chapter 6. The framework shows the relationship between patient attributes, system factors, and early diagnosis indicators, based on the relevant literature. The conceptual framework is then recreated with data items actually available for cancer patients. This is used to decide which case-mix factors should be adjusted for in the substantive analysis of early diagnosis trends (**Objective 2**).

In Chapter 7 a combined literature review and data analysis study on the association between different early diagnosis indicators and survival is presented, in the published paper *Which indicators of early cancer diagnosis are associated with short-term mortality and survival?*. The study is accompanied with further consideration of the results, leading to selection of stage as the indicator for the substantive analysis (**Objective 3**).

Statistical approaches for case-mix adjustment are reviewed in Chapter 8. Inclusion of the case-mix variables as covariables in a generalised linear model is identified as one attractive technique (**Objective 4**). Sampling variability and multiple comparisons are discussed as challenges for any

evaluation of geographic inequalities in Chapter 9. Multilevel modelling is introduced as one way to compare total geographic inequalities between different time periods (**Objective 4**). Multiple imputation, which uses auxiliary information on patients to impute their stage where it is missing, is then introduced in Chapter 10 (**Objective 5**).

In Chapter 11 the substantive analysis of the thesis, in which multilevel modelling and multiple imputation are used to assess stage trends and changes in geographic inequalities during 2008-2013, is presented in the second published thesis paper *Temporal and geographic changes in stage at diagnosis in England during 2008-2013: A population-based study of colorectal, lung, and ovarian cancers* (**Objective 6**).

Chapter 12 contains an evaluation of the impact of including different parameters in multiple imputations models; simulation to validate multiple imputation in one plausible scenario; and a sensitivity analysis measuring the error that would need to be present for conclusions of the substantive analysis to change (**Objective 7**). These analyses are presented together in the final (submitted) thesis paper *Multiple imputation to minimise bias from missing stage information in estimates of early cancer diagnosis in England: a population-based study.*

The discussion, Chapter 13, considers the future of early diagnosis monitoring. The need for multiple imputation for robust evaluations of stage is considered, with reference to the improvements in stage completeness and the changing meaning of missing stage data from 2008. Immediate potential applications of methods used in this thesis are discussed, including to assess the impact of the COVID-19 pandemic on early diagnosis (**Objective 8**).

2. Introduction: health services in England, 1990-2019

This Chapter provides an overview of how health services are organised in England and how this has changed from 1990. The different actors in cancer policy are introduced, and specific policies which may have affected early diagnosis are discussed. These details provide context for the analysis of early diagnosis trends during 2008-2013 which is reported in this thesis, and aid interpretation the changes observed. They are also used to identify the most appropriate level of resolution for the analysis of geographic early diagnosis inequalities.

2.1 History of National Health Service reform

The funding model for health care in the United Kingdom (UK) has changed little since the establishment of the National Health Service (NHS) in 1948. Since then, this organisation has provided acute, primary, and secondary care, free at the point of use to the public through single-payer funding from general taxation. The NHS dominates provision of healthcare in the UK: in 2014, up to 90% of the population relied solely on it for their care [6], and public expenditure (primarily through the NHS) accounted for 79.5% of the total health spending nationally [7]. Whilst the funding model has remained constant, the model for commissioning of local services has changed frequently, with numerous restructures between 1990 and 2013 (Figure 2.1).

Before 1991, both arrangement and delivery of NHS services in each local government ("local authority") area were managed by a local Health Authority. The National Health and Community Care Act 1990 split the "purchaser" and "provider" functions of these Health Authorities [8]. Following implementation of the Act in 1991, the 205 authorities purchased services from newly-formed Trusts (existing hospitals and care providers reformed into independent organisations) [9]. The 1990 Act was designed to create a marketplace for services to apply competitive pressure on the providers. The Act also initiated General Practitioner (GP) fundholding in some areas. Under this arrangement, consortia of family doctors took control of the local healthcare budget from their health authority, and had the choice to retain and reallocate any savings they could achieve.

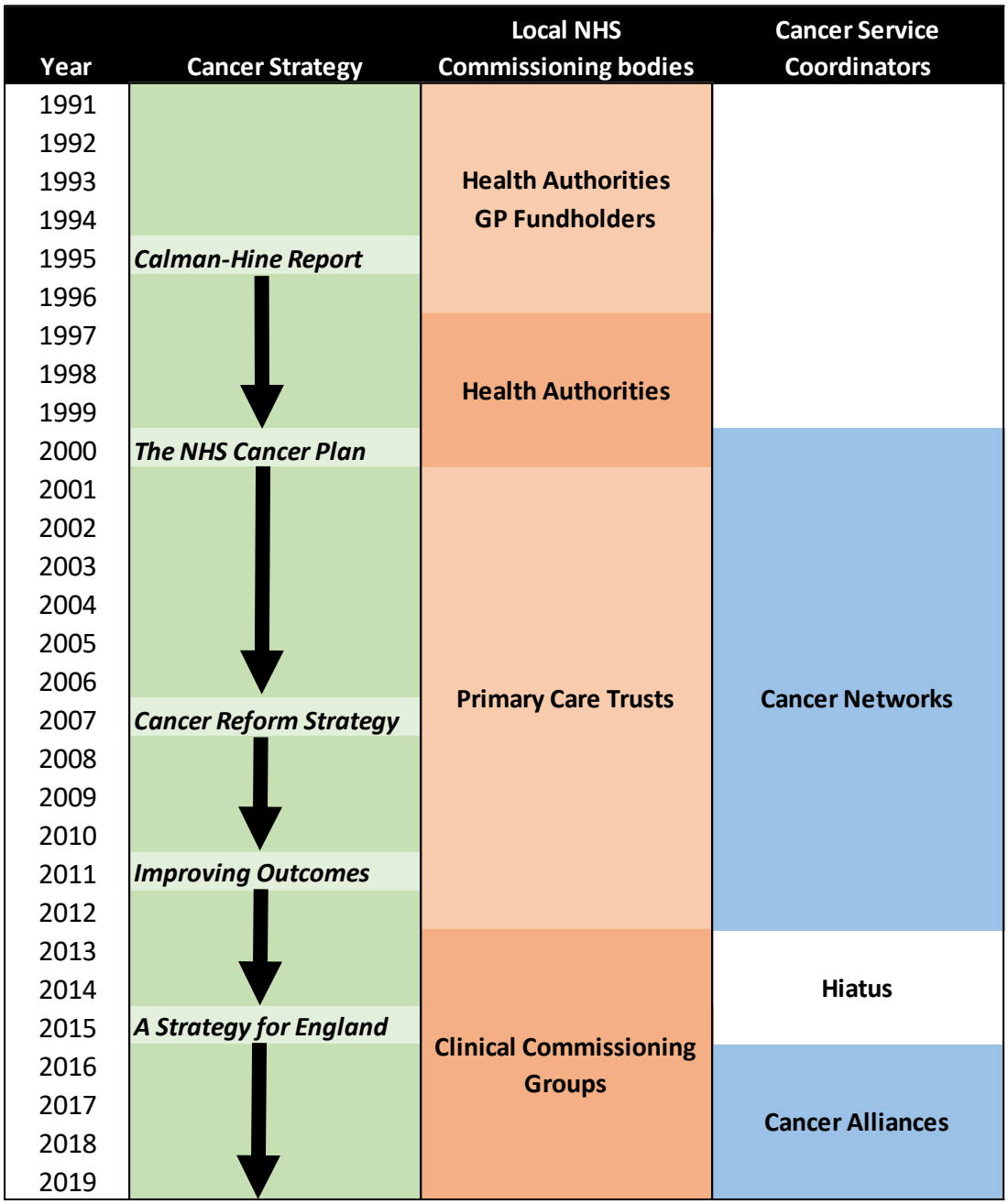
GP fundholding was abolished by the incoming Labour government in 1997. Commissioning responsibility reverted to Health Authorities in areas where it had been previously given to GP consortia. The Health Authorities also evolved around this time: in 1996 the 205 were reformed into 100 larger Authorities [10]. These took on additional responsibility for commissioning primary care, pharmacy, and dentistry services (which had previously been commissioned separately). In 2001 these 100 Authorities were replaced by 303 Primary Care Trusts (PCTs), which by 2006 had merged into 152 PCTs. The majority of these (~70%) had boundaries co-terminus with those of the historic Health Authorities. The net result was similar geographic boundaries for local commissioning in 1990 and in 2006, following several reorganisations. The fundamental divide between purchasers and providers introduced in 1991 persisted, however.

The Conservative Government elected in 2010 enacted a more substantial reform in 2013 through the Health and Social Care Act 2012 [11]. The Act had similar motives to the 1990 Act: to give family doctors control over NHS budgets and encourage competition amongst providers. The reorganisation dissolved the Primary Care Trusts (PCTs) and moved most of their responsibilities to newly-created Clinical Commissioning Groups (CCGs). CCGs are clinically-led organisations which include all the GPs in their geographic area. Most have territories co-terminus with the PCTs which preceded them, though some PCT areas were split between more than one new CCG. Each CCG buys secondary care services for patients on their GP practice lists and people within their territory who are not registered with a GP: on average 250,000 people per CCG in 2013 [3, 12]. The Act required CCGs to put services out to tender, allowing private companies to compete with established NHS Trusts to provide services. In practice, from 2013 spending through private organisations increased but remained small as a proportion of the total health budget (rising from 6.1% in 2013-14 to 7.3% in 2018-19 [13]). In addition to establishing CCGs, the Act moved local public health staff from PCTs to local authorities, and established a national public health agency, Pubic Health England (PHE).

The root-and-branch reform of the NHS was designed to improve services and make them more cost effective. Whatever its long-term impact, the implementation may have had an immediate effect on patient outcomes directly through interruption to the coordination of different NHS services, and indirectly through resources being diverted to effect the change. One policy evaluation of the reform reported that the complexity of the re-organisation and associated upheaval limited the effectiveness of the NHS in delivering services and responding to new challenges from 2013 [14]. Another study found evidence for disruption to cervical screening services due to the complexity of the reorganisation [15]. Otherwise, however, there is little empirical data for any effect of the Act on cancer service provision or outcomes.

In 2016 a new level of healthcare coordination above CCGs but below the national level was introduced: Sustainability and Transformation Plans (STPs). The 44 STPs are five-year plans covering all aspects of NHS spending for their region (with each STP encompassing 5 CCG areas on average). They do not replace existing organisations; each STP is a collaboration between the CCGs, NHS Trusts, and Local Authorities within its territory. The STPs were designed to integrate care across CCG boundaries, reduce variation in services and outcomes, and find efficiency savings. If an STP plan is approved by the Department of Health, NHS transformation funding is released to the constituent organisations from central government [16]. Organisations within an STP area are therefore financially incentivised to collaborate. This is to a large degree at odds with the philosophy of the 2012 Act, which sought to create efficiencies by encouraging competition – rather than cooperation – between service providers and between the CCGs themselves. This tension between integration of care and a structure designed to promote competition was finally resolved through the 2019 NHS Long Term Plan, which directed STPs to prioritise integrated care and removed the requirement for CCGs to automatically put services out to tender [17].

Figure 2.1 Timeline of national cancer plans during 1991-2019, against prevalent NHS local service commissioning bodies, and coordinators of cancer services.



2.2 National strategies for cancer control from 1995

The improvement of cancer services has been a focus of health policy in the UK, reflecting the large burden of the disease on public health and public concern about it. Evidence of lower cancer survival in England than in comparable affluent countries has provided a particular incentive to make improvements [18]. Policy has coalesced in successive national strategies for cancer coordinated by or written within the Department of Health, starting with the Calman-Hine Report in 1995 [19].

The Calman-Hine report, entitled *A Policy Framework for Commissioning Cancer Services*, helped establish integrated networks of care for cancer patients. The benefit of treating a cancer of a given type in a high-volume centre by a multi-disciplinary team (MDT) which specialises in treating that type has been established by clinical consensus and is frequently cited in health policy documents[1] [2, 21]. However, there are frequently insufficient volumes of patients with rarer cancers or atypical cases of common cancers in a local commissioner area to justify maintaining a specialist MDT for that type. It is thus often more practical for the MDT to serve several commissioners. Through covering a larger population, this arrangement achieves sufficient patient volumes to be financially viable and maintain its specialism. The Calman-Hine report recommended that patients requiring specialist treatment should have access to it in designated cancer centres, via referral from their local hospital [19]. Consequently, in 2000, 28 regional Cancer Networks were established to coordinate the flow of cancer patients to specialised MDTs and other services across commissioner territories [22]. The report also advocated good communication between providers, from GPs to Cancer Centres, and for robust cancer intelligence gained through national cancer registration.

The next major strategy, The *NHS Cancer Plan*, was published in 2000 and set forth a broad agenda for prevention, detection, and treatment [22]. The plan introduced waiting times targets to encourage providers to expedite symptomatic diagnosis, including a maximum two-week wait (2WW) to see a cancer specialist following general practitioner (GP) referral with possible cancer symptoms, and a 62-day wait to start treatment for certain cancers. It proposed greater use of specialised MDTs, and increased training of radiologists, oncologists, and surgeons. An additional £570 million was promised for cancer by 2003-04, representing a step-change in funding for cancer services. The plan also established a programme of work for the newly-established National Institute of Clinical Excellence (NICE) to evaluate cancer drugs and publish referral guidelines for suspected cancer for GPs. The new referral guidelines were published in 2005 [23].

The *Cancer Reform Strategy* was published in 2007. It expanded initiatives introduced through the 2000 strategy and introduced several more. The 2WW target for suspected cancer referrals was extended to include patients initially presenting at hospital or through screening, and the treatment waiting times target was extended to all cancers [2]. Screening programmes were also expanded. The coverage of the breast screening programme (first established in 1988) was expanded from women aged 50-70 to those aged 47-73 years. The Strategy set the target that the newly established Faecal Occult Blood Testing (FOBT) bowel screening programme should achieve national coverage for all men and women aged 60-75 by 2010. It established the National Awareness and Early Diagnosis Initiative (NAEDI) to raise awareness of cancer symptoms, and proposed initiatives to better understand the reason for international and deprivation differences in cancer outcomes.

[1] Quantitative evidence of the benefit from MDTs is more sparse, partially due the ethical and practical obstacles to robust evaluations. Further discussion in: [20] C.N. Prabhakar, K.M. Fong, M.D. Peake, D.C. Lam, D.J. Barnes, The effectiveness of lung cancer MDT and the role of respiratory physicians, Respirology 20(6) (2015) 884-8.

A fourth plan, Improving Outcomes: A Strategy for Cancer, was released in 2011. It sought to build on the 2000 and 2007 strategies, but emphasised the need to improve outcomes within an overall context of financial austerity, following the 2007/08 financial crisis [24]. There was an emphasis on patient choice and experience, and definition of the roles of different soon-to-be-established organisations created through the Health and Social Care Act 2012. The plan highlighted the need for symptom awareness campaigns, and later that year the 'Be Clear on Cancer' campaign was launched to raise awareness of common cancer symptoms and encourage people with them to go to their GP [25]. In contrast to previous and future plans, there was far less emphasis on target setting and evaluation of patient outcomes, and more emphasis on processes and efficiencies needed to maintain service levels with fewer resources.

The change in focus of the 2011 plan reflected wider financial constraints on public services from 2010. By 2013 growth in healthcare spending had slowed and the NHS was facing growing demand from an ageing population [4, 26]. The average wait for a GP appointment increased up to 2013, as did the average waiting time in A&E departments, reflecting the increased pressure imposed by austerity on these two services [27, 28]. As up to 75% of cancer patients first present in a GP practice or A&E department [29], it is possible that the increased pressure on these services reduced the quality or timeliness of diagnostic investigations for patients during this period, with a consequent reduction in early diagnosis. However, any impact may have been offset by increased diagnoses through screening, greater symptom awareness, new technologies, and organisational reforms to improve patient pathways. Updated NICE guidance published in 2015 explicitly advised more referrals to potentially detect cancer early, even at the cost of more tests on patients without cancer: GPs were advised to refer patients if their symptoms had greater than a 3% positive predictive value for cancer [30], contrasting with the 10% positive predictive value for cancer associated with referral through the TWW pathway prior up to 2007 [2].

The next major document in cancer policy was the report *Achieving World-Class Cancer Outcomes*: *A Strategy for England 2015-2020*. Published in 2015, the report was unique in being prepared by an independent cancer taskforce and not directly by the Department of Health (DH) as previous reports had been. Though ultimately commissioned by NHS England and other bodies funded by DH, the taskforce was chaired by the Chief Executive of the charity Cancer Research UK, Sir Harpal Kumar, and included members from DH and arms-length bodies of DH; Royal Colleges; charities; and NHS trusts [3]. Amongst other measures, the plan recommended improving access to diagnostic tests for GPs, with the intention of expediting time to definitive diagnosis or exclusion of cancer; investments to address bottlenecks in diagnosis due to resource shortages in radiology, radiography, and endoscopy; and a focus on patient experience. One feature, radically different to the 2011 plan, was the setting of specific targets covering prevention, early diagnosis, patient experience, and survival. Accompanying these targets were plans to introduce surveillance of outcomes at the national, Cancer Alliance, and CCG level. Implicit in the refocus was the idea that target-setting and performance measurement would act as drivers for improvement beyond 2015.

2.3 Organisation of cancer services from 1995

Responsibility for cancer has resided with local healthcare commissioners throughout 1990-2016: first with Health Authorities (1990-2001), then PCTs (2001-2013), and then CCGs (2013-present). Their territories have covered populations of on average 250,000-350,000 people each.

Following the recommendations of the Calman-Hine report, from 2000 to 2013 28 Cancer Networks coordinated the flow of cancer patients to specialised MDTs and other services across PCTs in their territories [31]. At the time of the dissolution of PCTs in 2013 the Networks were also disestablished, however. In 2015 a commitment was made to establish new Cancer Alliances, which would be tasked with driving service improvement and coordinating patient pathways between CCGs within their territories [3], fulfilling an equivalent role to the Cancer Networks. From September 2016 19 Alliances and Vanguards[2] were established, each covering on average 11 CCGs and 2.8 million people [32, 33]. There was therefore active coordination of cancer services at a level above the local commissioner during 2000-2016, though without formal structures for coordination between April 2013 to September 2016 (figure 2.1).

In addition to CCGs and Cancer Alliances, there is specialist commissioning of specific national services. These services include the NHS Breast Screening Programme (from 1998 [34]), the NHS Bowel Screening Programme (from 2007 [35]), and the 'Be Clear on Cancer' symptom awareness campaigns (from 2011 [25]). From 2013 these national services have been commissioned directly by Public Health England. Prior to that they were commissioned by the Department of Health.

2.4 Summary

The period 1990-2019 was defined by transition for the health services in England. From 1995 successive national strategies for cancer were produced, each aiming to achieve a step-change improvement in outcomes through better diagnosis and treatment. The changes effected at the time of the strategies occurred against a backdrop of wider organisational change. There was one large health service reorganisation in 2013, and smaller interventions before and after. From 2000, per capita spend on cancer rose sharply, but it plateaued from 2010 as financial austerity set in [4]. These systemic changes are also relevant to interpreting early diagnosis trends during 2008-2013.

An assessment of national trends in early cancer diagnosis will speak to the combined effect of all these changes during 2008-2013. Additionally, the success of interventions for specific cancer types in this period, such as bowel cancer screening, can be evaluated by assessment of early diagnosis trends for those types.

Up until April 2013 PCTs had responsibility for cancers services within their geographic territories. Thereafter, CCGs were responsible. To ensure the relevance of the results, the current commissioner

[2] In Greater Manchester and two areas in London, Cancer Vanguards instead of Alliances fulfil the functions of Cancer Alliances. The Vanguard areas are empowered to pilot new models of care and integration of services within their areas. Hereafter in this document "Alliances" refers to Alliances and Vanguards.

territories at time of the analysis – CCG areas – will be used in this thesis for the analyses of geographic inequalities.

3. Performance measurement and target-setting in health policy

One concern about the single-payer structure of the NHS is that there is a lack of pressure to perform exerted by market forces. Pressure might be expected in a setting where customers have a chance to select for cost effectiveness in an open market. However, obstacles to well-functioning marketplaces for healthcare with any existing model have been documented [36], due to customers lacking the necessary expertise to discriminate good from bad services, and constraint of choice by geographic practicalities [37]. For the NHS, central government provides this external pressure.

Approaches taken by the government to create pressure to perform fall broadly into the following categories: i) competition to provide services ii) clinical governance iii) performance measurement and target setting. Each of these potentially had some positive impact on cancer outcomes during 2008-2013. Competition has been encouraged through the purchaser and provider split of 1990, and from April 2013 through requiring the purchase of services to go out to tender. Clinical governance has been a statutory duty for NHS providers from 1998, and encompasses a variety of behaviours and practices designed to encourage clinical excellence, risk management, and patient centred care. The work of this thesis is closely aligned with the final mechanism: performance measurement and target setting. Targets became increasingly used in the NHS from 2000, being embedded in the NHS Plan (2000) and the Public Service Agreements which attached specific goals to Treasury spending commitments. The government sought to use these to ensure that a step-change increase in NHS funding from 2000 was matched by effective use of the increased resources.

This Chapter considers the strengths and limitations of performance measurement and target setting for the NHS, and then specifically how this thesis can contribute to performance measurement of early cancer diagnosis.

3.1 Performance measurement to improve local services

There has been considerable debate about the impact of target-setting in the NHS, with findings of both positive and negative effects [38, 39]. The increased use of targets for hospitals and commissioners from 2000 has been criticised for creating a culture of passivity and risk aversion, lowering morale, and sometimes resulting in gaming and neglect of services not covered by the targets [40-42]. However, a review of the NHS reforms by the King's Fund also found that target setting was the only performance improvement mechanism for which there was clear evidence of effectiveness. That review found sparse evidence for other approaches, such as inspection and competition between providers, for service improvement [37]. This is potentially to be expected if targets focus efforts on providing empirical evidence of improvement (i.e. hitting the target at any cost), potentially at the expense of the underlying change envisioned by the target-setter, but not entirely captured by the target.

Policy analysts have also raised concerns about the quality of intelligence gained by evaluating performance against targets. Target-setting has been assessed as suitable to identify failing providers where a targeted intervention is required, but inadequate by itself to lead the transformation of services from adequate to good or excellent [43]. Bevan *et al* (2006) describe measures of healthcare performance against targets thus:

"They are tin openers rather than dials […] they do not give answers but prompt investigation and inquiry, and by themselves provide an incomplete and inaccurate picture." [42]

Statisticians have also questioned the use of performance measures derived from sparse data, where the random play of chance may have more impact than an institution's performance in determining its observed results [44]. They have also raised concerns that failure to adjust for case mix: the population composition which affects health outcomes independently of service performance. This may lead the services which accept more challenging patients to return poorer results and so be penalised rather than supported to improve, since their outcomes aren't reflective of their underlying performance.

Target setting has therefore been found to be effective at ensuring specific service levels are achieved. However, measurement of performance against targets also carries risks of demoralising staff, encouraging unintended behaviours contrary to the objective behind the target, and the potential for misleading and uninformative results which penalise services providing good or excellent medical care.

3.2 Performance measurement for overall evaluation of health service functioning

Performance against targets is used to assess the functioning of health services at a national level by the media, public, and politicians in England. Perhaps the most prominent target used for this during 2008-2013 was the four-hour A&E waiting times target. It states that 95% of patients attending A&E should be admitted, transferred, or discharged within four hours. It has been in place since 2004, and is hugely important in public and political discourse as a measure of NHS functioning [27]. A pledge to meet the target has been integrated into the handbook to the NHS constitution [45], and performance against the target remains a mainstay of media reporting on the NHS.

Though often taken as a measure of time spent waiting for consultation in A&E, the 4-hour target is more often an indicator of lack of capacity in inpatient wards [46]. Non-urgent A&E attendances are typically assessed, treated if needed, and sent home promptly. Patients with severe or complex illnesses will need to be admitted, but they may have to remain in A&E for some time if beds aren't available in inpatient wards. Bed shortages can occur when current inpatients can't be discharged promptly, for example if there isn't capacity in local social care services to receive them, or if there aren't resources to promptly do all tests needed before they can be safely discharged.

Performance against the 4-hour target can therefore be a systemic health check for the whole health and care system, locally and nationally. The target aids public understanding of how services are functioning at times of increased pressure (e.g. during winter flu season or other times of heightened activity), and informs the wider discourse about the performance of the NHS.

Cancer survival statistics have played a similar role in shaping the debate about NHS cancer services. Like the A&E waiting times target, they reflect several different aspects of patient care: screening, diagnosis, treatment, and long-term follow up; and hence provide a systemic indicator for the functioning of services. Evidence of the lower survival in England provided a particularly important stimulus for increased investment in and reform of cancer services from 2000. It was cited as a key reason for the changes introduced through the national cancer plans from 2000 [22] to 2015 [3]. It was also a key justification for the 2013 NHS reorganisation. The factsheet supporting the 2012 Health and Social Care Act which effected the reorganisation cited England's lower survival in the "case for change" as follows:

> *"**Need for improvement**. At its best, the NHS is world-leading, but there are important areas where the NHS falls behind those of other major European countries. If we had cancer survival rates at the average in Europe, we would save 5,000 lives a year."* [11]

The prominence of cancer survival statistics in these key policy documents show their power in stimulating efforts to improve services. It also shows comprehension of their importance in the media and in public opinion. That the survival statistics have at different times been cited as justifications for quite different reforms (for example, policies of competition effected through the Health and Social Care Act 2012, and policies of integration in the 2019 Long Term Plan) reflects both differences in opinion about the best approach for service improvement, and uncertainty about the reason for the survival deficit between the UK and equally developed nations.

Academic researchers have sought to understand the reasons for international survival differences, for example whether they may be due to differences in stage at diagnosis or access to surgical treatment [47, 48]. In 2009 the Department of Health initiated the International Cancer Benchmarking Partnership (ICBP), a collaboration of academics and health professionals from six high income countries tasked with understanding the differences in survival between their countries [49]. The rationale for the ICBP is that understanding of the reasons for differences will stimulate improvements in each participant country and raise survival towards the best achievable. A recent (2019) analysis from the partnership reported persistent differences in survival between the participating countries, and hypothesised that stage differences are likely to play a role in these [50].

3.3 The current landscape of cancer intelligence and cancer target setting in England

The foundation of cancer intelligence in England is cancer registration, the collation of population-based data on new diagnoses. Registration in England dates back to 1929, when it was initiated by

the Radium Commission. It is currently the responsibility of National Cancer Registration and Analysis Service (NCRAS), which has legal permission (through the NHS Act 2006) to collate and report information on all cancers diagnosed nationally [51]. The Act allows collection of data on patients without their consent, though it is possible to actively opt out as cancer is not a notifiable disease [52]. NCRAS aims to collect data on all cancers diagnosed in the NHS, and also receives data on some patients treated privately [53].

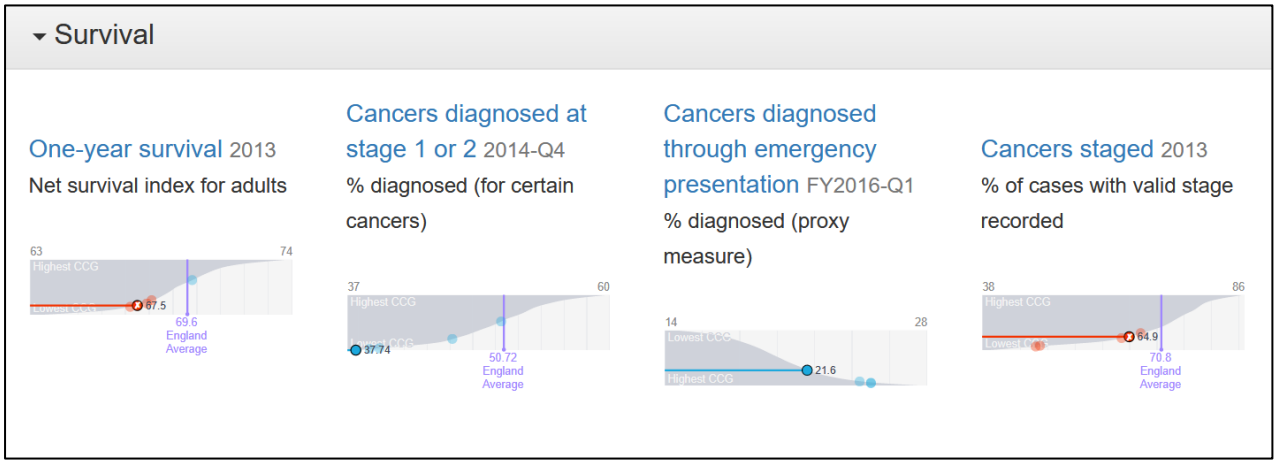
Cancer registrations at a minimum include identity and demographic information on the patient, the topographical site the tumour originated in, and the date of diagnosis. This information is then linked the NHS Central Register for the patient's date of death, and used to calculate survival time from diagnosis. Information on stage and morphology may also be collated in registrations. Historically this information has been very poorly completed, but there have been considerable improvements in recent years: for example, in 2001 fewer than a quarter of malignancies had information on the stage at diagnosis recorded, but by 2018 recording was over 85% [5]. Furthermore, in recent years it has become possible to link registration records to information from other sources, for example on deprivation and activity in secondary care around the time of diagnosis [54].

As discussed in section 2.2, the 2015 *Strategy for Cancer* set specific targets for cancer outcomes, and recommended the introduction of surveillance to monitor progress against these at the national and local commissioner level [3]. The evidence for a survival gap between England and other developed countries, and for geographic inequalities in early diagnosis and other outcomes, was a key motivation for this. The Strategy set the goal of raising survival from cancer in every CCG towards the highest achieving CCG, so that one-year cancer survival would reach 75% nationally by 2020 [3]. Some other specific national objectives set for 2020 included:

- An increase of 10-year net survival to 57%
- A reduction in between-CCG variation in one-year survival;
- A reduction in the gap in cancer survival between older and younger patients;
- 95% of patients to have a diagnosis or cancer excluded within 4 weeks follow GP referral;
- Percentage of cancers diagnosed at stages I or II (of all with stage recorded) to reach 62%;
- The percentage of cancers with stage recorded to increase;
- 75% uptake of Faecal Immunochemical Test (FIT) screening for bowel cancer;
- 96% of cancer patients waiting no more than 31 days to start treatment
- 85% of patients waiting no longer than 62 days total from referral to treatment.

Implementation of the surveillance programme took the form of a dashboard, which is run by PHE and went online as the public-facing *CancerData* website in 2016 [5]. The dashboard presents quarterly statistics on early-stage diagnosis and emergency presentation, and annual statistics on survival, incidence, and patient experience. CCG-level late diagnosis percentages which are statistically significantly different (at the 5% level) from the national average are flagged as being either significantly higher or lower (Figure 3.1).

Figure 3.1 Presentation of early diagnosis statistics on PHE's *CancerData* dashboard, showing the national average results (in purple) compared to results for a specific CCG (in blue and red) [5].



In addition to surveillance initiated after 2015, targets have also been integrated into payment-for-performance schemes for CCGs. For example, in the 2016/17 financial year CCGs were rewarded for either achieving a 4 percentage point increase in recorded diagnoses at stages I or II, or achieving overall 60% early-stage diagnosis (from an a non-case-mix adjusted group of common cancers) [55].

There are reasons for concern about the approach initiated through the 2015 Strategy for Cancer. The targets set for screening uptake, and waits for diagnosis and treatment, were the most likely to be practically achievable, as they are operational measures. By contrast, early-stage diagnosis and survival improvements are more challenging to effect; these can require significant investment to change normal management of patients, breakthroughs in technology for diagnosis and treatment, changes in behaviour and circumstances in the population, or some combination of all these. The national and CCG-level survival targets set in 2015 were not accompanied with an explicit justification or roadmap, however. Instead, highest previously observed results for any CCG were used as benchmarks or targets for the other CCGs to meet. This approach is also vulnerable to criticism. It implicitly assumes that variation is avoidable; that the CCGs with the best observed outcomes previously achieved them through good practice; and that other CCGs could replicate their results. Most of these assumptions are likely to be at least partially incorrect. Case-mix factors (including composition of different cancer types) and random chance can have a strong influence on results, particularly if patient numbers are small [44, 56] (further discussion in Chapter 9). Though there was no penalty for CCGs which failed to achieve the targets, there is still a risk they were not practically achievable with the resources available, or that performance against them was determined by factors outside a commissioner's control. In that case, they may have distracted from other service improvement efforts, or caused other counter-productive behaviours.

3.4 Is a new evaluation of early diagnosis needed?

The impact of cancer survival statistics in stimulating action to improve services has been discussed in sections 2.2, 3.2, and 3.3. As with performance against the 4-hour A&E waiting times target, monitoring of cancer survival has provided a systemic measure of performances of cancer services for health services and the public, allowing comparisons to other countries and between time periods.

Early diagnosis statistics have the potential to play a similar role. They have a distinct advantage over survival statistics in potentially being more rapidly available, as long-term follow up is not required. They also provide a specific measure of performance in areas relevant for early diagnosis: awareness of symptoms, health-seeking behaviour, available screening technology and uptake, and prompt referral and diagnosis. Early diagnosis statistics are particularly relevant for the period 2008-2013, when several interventions were being implemented at a time of organisational change, and when cancer intelligence was improving.

There are, however, reasons for caution in the use of these statistics, as discussed in sections 3.1 and 3.3. To be useful the statistics must be interpretable and, as much as possible, reflect the actual performance of health services. This requires consideration of factors such as patient case mix, use of an appropriate early diagnosis indicator, and approaches for handling of data which are sparse, of poor quality, or missing.

It is notable that an analysis of national trends in stage or changes in geographic inequalities during the period 2008-2013 has not previously been conducted. PHE report the percentage of patients diagnosed at stage I or II from a selection of 10 common cancers. However, that statistic is not case-mix adjusted, and generated only from patients whose stage was ascertained by the cancer registry (approximately 75% of patients in 2013 [57]), and so may give an incomplete account of changes in the whole population. There is therefore an opportunity for a new contribution in this area – so long as the methodological challenges outlined above can be addressed convincingly.

4. Clinical context to the cancers analysed in the thesis

4.1 Introduction

Early diagnosis is evaluated separately for three cancers in this thesis: colorectal cancer, non-small cell lung cancer, and ovarian cancer. This is to allow in-depth analysis of changes for a range of malignancies. As different cancers differ in typical presentation and patient profiles, and technology available for screening, diagnosis, and treatment, it is expected that trends in early diagnosis will be different for them. For example, the bowel screening programme rolled out nationwide during 2006-2009 would only have directly increased early diagnosis for colorectal cancer, although it is possible it would have improved cancer awareness more generally as a positive side effect.

The cancers chosen for analysis have been selected because of their relatively high incidence and unfavourable average stage at diagnosis, emergency presentation risk, and survival at the start of the analysis period in 2008. There was therefore considerable scope to improve early diagnosis for these malignancies during 2008-2013, and any increases could result in survival improvements at the population level.

The clinical context to three cancers considered is described in this Chapter. Details of initial presentation and current practices for diagnosis and treatment are discussed, with a particular emphasis on changes during the period 2008-2013. This contextual information informs a conceptual framework for the patient and system factors which determine early cancer diagnosis, and is used to interpret the trends in early diagnosis which are estimated.

4.2 Colorectal cancer

Colorectal cancer is one of the most common cancers diagnosed, with approximately 35,000 new diagnoses registered in England every year [58]. Symptoms can include bleeding from the anus, changes in bowel habit, a lump which can be felt inside the anus, pain in the lower abdomen, tiredness, weight loss, and vomiting [59]. One or more of these symptoms may prompt urgent GP referral or screening tests.

In England, screening may be done at home with either guaiac Faecal Occult Blood Tests (gFOBT) or with Faecal Immunochemical Tests (FIT) [35, 60]. Both test for blood in faeces, but the FIT is more specific as it detects human haemoglobin, whilst the gFOBT can test positive from animal blood ingested as part of a normal diet (for example, from red meat). A positive test can indicate bleeding within the GI tract which may be due to cancer, with the FIT having sensitivity for cancer of around 80% or higher depending on the false-positive tolerance chosen [61]. Patients with a positive test will be referred for endoscopy. Biennial screening with gFOBT at home for colorectal cancer was introduced in 2006 for persons aged 60-69 in England, with national coverage achieved by 2009 [60]. In 2010 the programme was extended to people aged 70-74 [62]. In June 2019 gFOBT was replaced with FIT nationally [63]. Geographic variation in screening uptake has been documented, partly aligned with geographic deprivation differences [64].

In 2013 the bowel screening programme was further extended to include a one-off flexible sigmoidoscopy investigation at age 55 for all men and women in addition to the biennial gFOBT for adults aged 60-74 [62], following evidence from a large RCT that sigmoidoscopy conferred reductions in both cancer incidence and mortality [65]. Sigmoidoscopy is used for the examination of the lower part of the large bowel and removal of polyps (pre-cancerous growths) in order to stop cancer developing. It is also effective for the diagnosis for patients with certain combinations of symptoms, detecting 95% of cancers in patients who don't have iron deficiency anaemia or an abdominal mass but who do have other symptoms of colorectal cancer [66].

Patients with either specific symptoms or a positive result from a screening test may be referred for sigmoidoscopy or colonoscopy for diagnostic investigation. Which investigation is chosen partly depends on what the presentation indicates about where the tumour is located: for fresh rectal bleeding sigmoidoscopy is indicated, but for anaemia or altered bowel habit a colonoscopy is needed as cancer is likely to be higher in the bowel than sigmoidoscopy can reach. Biopsies may be taken as part of both sigmoidoscopy and colonoscopy investigations. In recent years CT colonography has become available, which performs the same function as colonoscopy but is not invasive, and is optimal for patients who are frail or otherwise unable to tolerate colonoscopy [62]. Cancer stage is established on the basis of data collected from one or more of CT, PET-CT, or MRI scans, or with ultrasound for rectal cancer, and the pathology from any biopsies taken [35].

Treatment involves surgical removal of the main tumour or tumours and affected tissues. Preoperative or postoperative radiotherapy or chemotherapy can be used to reduce the risk of recurrence for patients with advanced disease [47]. In 2008 in England, around 75% of colorectal patients had surgical treatment, but only 60% had a major resection associated with complete removal [67]. Total mesorectal excision is the optimal treatment for rectal cancer, and became common in England in the late 2000s [68]. Preoperative radiotherapy has been found to confer a survival advantage for rectal cancer [69], however, NICE guidelines ultimately leave treatment decisions up to the multidisciplinary team (MDT). Studies have shown there is large variation between NHS Trusts in the use of radiotherapy for rectal cancer [70]. Adjuvant chemotherapy is typically used for stage III colorectal cancer.

Tumours which are not detected whilst still small in size may grow to obstruct the bowel. The blockage will cause vomiting and constipation, and will be fatal if the patients do not have emergency surgery or a stent inserted to unblock the gastrointestinal tract. Not all patients are fit for surgery, and for the patients for who are, time is needed to prepare for it. Stents to unblock the bowel can be crucial in allowing a delay before surgery or as an alternative to it in cases where the patient can't be operated on, and these have become increasingly used since they were first introduced in the 1990s [71, 72]. Stenting itself carries risks of the stent moving from where it was placed or perforating the tumour [73], and emergency surgery when stenting can't be done is also associated with high mortality, as a major surgery on patients who are already acutely unwell [74].

Overall net survival from colorectal cancer was lower in England than in other affluent countries through the period 2000-2009. There is, however, evidence of a step change between the period 2005-09 and 2010-12, concurrent with the national interventions to increase early diagnosis and expansion of bowel cancer screening (see Figure 6.1 [75]).

In summary, the period 2008-2013 was one of considerable change in colorectal cancer screening and diagnosis. The changes may have led to improvements in the proportion of patients diagnosed early. Improvements in treatment around this time may also have raised survival at each stage during 2008-2013. There is evidence for geographic inequalities both in screening uptake and use of different treatment options such as radiotherapy.

Figure 4.1 Trends in 1-year and 5-year net survival form six common cancers during 1995-2013, taken from Walters, Benitez-Majano, Muller, *et al* (2015) [75], showing large increases in colorectal, lung, and ovarian cancer survival during the period 2008-2013.

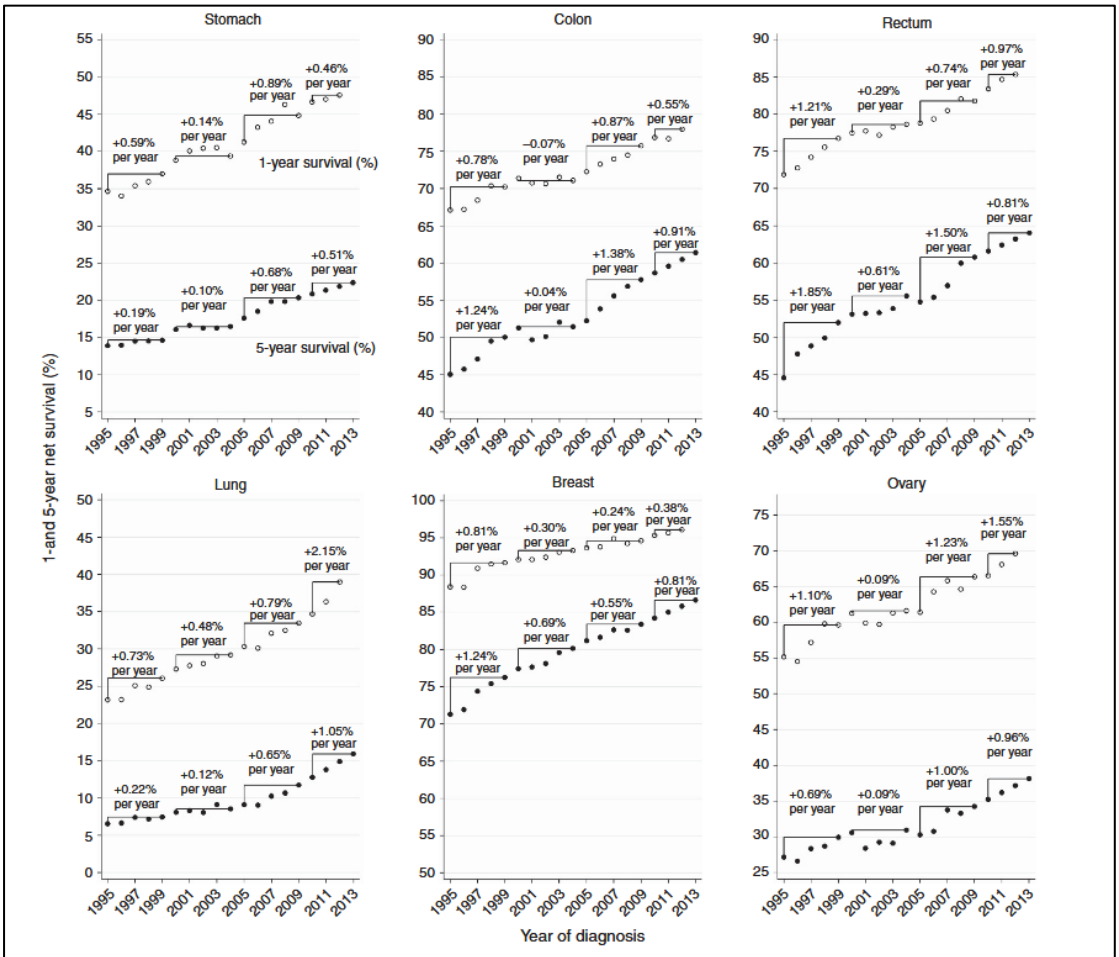


Figure 3. Trends in 1- and 5-year net survival in England by year of diagnosis. Unfilled and filled markers represent estimates of 1- and 5-year net survival, respectively. The average arithmetic improvement in survival between years is shown for the calendar periods 1995–1999, 2000–2004, 2005–2009 and 2010–2013 above the trends.

4.3 Non-small cell lung cancer

There are approximately 37,000 new cases of lung cancer registered in England every year [58]. Lung cancer encompasses different morphological subtypes associated with different risk factors and tumour grade (speed of tumour growth). Specifically, small cell lung cancer is overwhelmingly caused by smoking, and mesothelioma of the lining of the lungs is strongly associated with asbestos exposure [76]. Both of these cancers are typically diagnosed at higher grade, with more aggressive growth and spread compared to other lung cancers, and at late stage. The incidence of mesothelioma and small cell lung cancer has decreased during 2008-2013, tracking historical changes in risk factor exposure [77, 78]. To avoid any confounding effect from these changes in the estimates of early diagnosis trends conducted in this thesis, and to simplify the interpretation, these cancers were excluded and only non-small cell lung cancer (NSCLC) was evaluated. NSCLC includes morphological types that are more similar in their biological behaviour; together these comprise 85% of lung cancers diagnosed every year.

The first symptoms of lung cancer can include a new cough, a change in an existing cough, recurrent chest infections, pain, breathlessness, coughing up blood, tiredness or lack of energy, and unexplained weight loss [79]. Many of these symptoms, such as breathlessness and tiredness, are very non-specific to lung cancer, particularly amongst the older smokers and ex-smokers who comprise the majority of new cases. As a result, early symptoms may be missed in GP consultations [80]. A pilot “Be Clear on Cancer” symptom awareness advertising campaign ran on television in the central TV region (East and West Midlands) in Autumn 2011, with national poster campaigns following in spring 2012 and summer 2013 (Figure 6.2). The campaigns were accompanied by increased symptom awareness, GP activity, and referrals [81], but empirical evidence of a stage shift was not reported. No screening programme existed for lung cancer during 2008-2013, though a pilot randomised controlled trial (RCT) of low-dose computerized tomography (CT) screening conducted during 2011-2013 indicated the introduction of one may confer mortality reductions [82], and a commitment was made to extend screening in the 2018 NHS Long Term Plan [83].

Figure 4.2 "Be Clear on Cancer" symptom campaign poster highlighting breathlessness as a symptom of lung cancer [84].

Where cancer is suspected, GPs may make direct referrals for a chest X-ray to test for lung cancer [85]. A normal X-ray does not exclude lung cancer, but the majority of symptomatic patients will have an abnormal chest X-ray. If the X-ray is abnormal or symptoms are highly specific to lung cancer, a CT-scan will be done to diagnose the cancer.

For patients who are potentially suitable for curative surgery a PET-scan is also done for full staging, and a brain MRI is done to check for brain metastases. Biopsies can be taken via bronchoscopy if tumours cells are located in the central part of the chest, and more recently endobronchial ultrasound scans (EBUS) have been introduced, which can perform an ultrasound of the airways and also take biopsies from lymph nodes. In 2011 NICE recommended that every Cancer Network should have at least 1 centre with EBUS or equivalent technology [85].

Treatment with curative intent involves surgery to remove the cancer, or radiotherapy [85]. Chemotherapy may be offered to patients diagnosed at stages II or III not undergoing surgery, or used alongside surgery and radiotherapy. However, the majority of patients present with inoperable later stage disease [68]. Within England, higher deprivation and rural geographic location are both associated with lower likelihood of aggressive treatment and lower survival [86, 87]. Evidence suggests larger, more specialised MDTs are more likely to consider patients suitable for radical surgery and have better outcomes, but they are less likely to accept referrals from outside their areas [88]. Some of the geographic variation in treatment access and uptake may be explained partly by pervasive fatalism amongst both patients and clinicians, with a tendency against recommending disruptive treatment schedules far away from a patient's home if they are frail and the prognosis is poor even with treatment.

Mortality following resection was historically high and a deterrent to surgical treatment, but had fallen to 3-6% by 2004-2010, a relatively low rate given the otherwise poor expected outcomes for patients not undergoing aggressive treatment [89]. Surgery was historically performed by cardiothoracic

surgeons with a mixed practice, but from the mid-2000s there has been a trend towards increasing numbers of specialised thoracic surgeons, who have higher volumes of lung cancer resections and improved outcomes [90].

The landscape of lung cancer screening and diagnosis did not change as much as for colorectal cancer during the period 2008-2013. However, whilst no large-scale screening programme existed at this time and there was no paradigm-shift in the technology available for diagnosis, the symptom awareness campaigns from 2011 may have had an impact on promptness of symptomatic presentation to the GP and onwards referral. Additionally, there were improvements in stage-specific treatments and outcomes in this period, which may have led to an attitude change with respect to the importance of early diagnosis amongst doctors and patients. Geographic inequalities with respect to treatment are well documented, but there is less data available on geographic inequalities in early diagnosis.

4.4 Ovarian cancer

There are approximately 6,000 new cases of ovarian cancer registered in England every year [58]. Similarly to lung cancer, the first presenting symptoms may be vague or non-specific, and women may have symptoms for several months before cancer is considered likely [91]. First signs can include bloating, loss of appetite, pain in the pelvic region, changes in bowel or urinary habits, and tiredness or weight loss [92]. In Spring 2013 local “Be Clear on Cancer” symptom awareness campaigns for ovarian cancer ran in selected Cancer Network areas in England, focusing on bloating as an important symptom which should prompt GP attendance. One evaluation found the campaign resulted in increased awareness of bloating as a potential symptom of the cancer, but no evidence was found for a difference in either GP attendances with symptoms or an increase in diagnoses following urgent referrals [81].

If a GP suspects ovarian cancer, they may request a CA125 test [93]. However, this is not highly specific to ovarian cancer: for pre-menopausal women referral would not normally be on the basis of raised CA125 alone. GPs also have direct access to ultrasound scans, and they may either utilise this or make a referral based on CA125 alone, with ultrasound scans performed later at the hospital. If the CA125 levels are raised and/or ultrasound shows problems with the ovaries, the GP will either refer the patient on the two-week wait pathway (23% of all patients in 2004-08 [54]) or they will be referred directly to a bowel surgeon, a general oncologist, or a gynaecological oncologist (20% of all patients in 2004-08). Late diagnosis is common, with over 30% of women diagnosed via emergency presentation (emergency GP referral or A&E attendance) in 2004-2008 [54].

If CA125, ultrasound, and the patient’s clinical condition indicate a malignancy, a CT scan of the pelvic and abdomen is performed to establish the extent of disease. At that point, the patient may undergo surgery immediately, neo-adjuvant treatment then surgery, or neo-adjuvant treatment alone [94, 95]. Depending on the hospital attended, the patient may either be seen by a specialist MDT (including a gynaecological oncologist, a medical oncologist, radiologist, pathologist, surgeons, and a

clinical nurse specialist) or a local MDT (which has a cancer unit lead and a gynaecological oncologist).

Over a third of women diagnosed with ovarian cancer die within one year, with risk factors for lower survival including older age and emergency presentation [96]. Due to the lack of specific symptoms at early stages, late-stage diagnosis is common, and is strongly associated with lower survival [97]. Low albumin levels at time of diagnosis are also associated with lower survival, and though the causal relationship is not clear, low albumin is also associated with poor nutrition, metabolic effects from the tumour, small bowel obstruction and build-up of fluid in the abdomen [98]. For patients who undergo major treatment, post-operative complications include haemorrhages or toxicity from chemotherapy [95]. Nationally, ovarian cancer survival was lower in England than in some other comparable affluent countries during the period 2008-2013 [75], although survival improved during this period and there was evidence of a step-change improvement from 2010 (Figure 6.1). These changes in survival have been attributed to increasing centralisation and specialisation for treatment [99], in particular increases in the percentage women having surgical treatment in specialist cancer centres or with specialist oncologists [100].

In summary, and similarly to lung cancer, the literature indicates that diagnosis and management of ovarian cancer remained the same during 2008-2013 in England, with late diagnosis being common. An evaluation of an ovarian cancer symptom awareness campaign run in 2013 did not return evidence for any change in promptness of patient presentation with symptoms or urgent referral for testing. However, there were increases in net survival nationally during this period, which have been attributed to improvements in treatment.

4.5 Summary

There were substantial changes in screening and diagnosis of colorectal cancer during 2008-2013, including the introduction then expansion of the gFOBT national screening programme and increasing use of flexible sigmoidoscopy. For non-small cell lung and ovarian cancers there were no such substantial changes. However, for non-small cell lung cancer there was limited evidence for a positive impact from a symptom awareness campaign on early diagnosis in 2011, as well as improvements in treatment during this period, which could have had a knock-on effect in changing attitudes on early diagnosis. Similarly, for ovarian cancer, there was no evidence for an immediate effect of a symptom awareness campaign on presentation or referrals, but some indication that improvements in treatment conferred survival improvements in this period.

As discussed in Chapter 2, the changes (or lack of changes) for management of the three cancers in this period occurred against a backdrop of increased investment to 2010; standardisation of diagnosis and treatment practices through National Institute of Clinical Excellence (NICE) guidelines (with guidance relevant to these cancers published in 2004, 2005, 2008, and 2011); and from 2000 the availability of the two-week wait referral pathway for use by GPs (and, from 2008, hospital doctors) to urgently refer patients with suspected cancer. These system-wide improvements may have had an

independent or multiplicative effect on early diagnosis outcomes in this period alongside disease-specific changes in diagnosis and treatment. Both the system-wide and cancer-specific changes need to be considered as factors which explain early diagnosis trends.

5. Analysis datasets and data items

This Chapter describes the different datasets and data items which are used to meet the thesis objectives. The datasets available include cancer registration datasets up to 2013 and other datasets derived from NHS disease-specific clinical audit datasets and Hospital Episode Statistics (HES) data up to that year. Together they contain data items on:

- Where and when the patient was diagnosed with cancer and biological characteristics of the tumour (topography, morphology);
- Patient's stage at diagnosis;
- Demographic characteristics of the patients;
- Contextual information about the patient's activity in the health system prior to diagnosis;
- Patient's subsequent treatment and survival.

The interpretation of the more complex items is discussed in this Chapter. Particular attention is given to issues of data quality and completeness, as these may present obstacles to a robust analysis of trends.

5.1 Data linkage strategy

Cancer registration data was used to identify patients diagnosed with one of the three cancers considered in this thesis during 2008-2013. Specifically, individual registrations of a first, invasive, primary neoplasm for all adults aged 15–99 diagnosed with either colorectal, non-small cell lung or ovarian cancer (ICD-10 codes C18-20, C21.8; C33-34; and C56-C57.7 [22]) in England during the period 2008-2013 were provided by Office for National Statistics (ONS).

The registrations form the 'spine' of all the data analysed; only patients with a registration were analysed, even if information on cancer in other patients was reported in other databases. Additional information on these patients from other sources was then joined to the registrations. The registration data were linked to the Cancer Analysis System (CAS) dataset provided by Public Health England using patient's NHS number and postal code; to datasets from the National Bowel Cancer Audit [101] and the National Lung Cancer Audit [102] (both current to 31 March 2013); to the Routes to Diagnosis dataset (current to 31 December 2013) [54], and to data derived from Hospital Episode Statistics (HES) current to 31 December 2013, by Adrian Turculet in the Cancer Survival Group. Registrations were also linked to the National Health Service Central Register to obtain information on the death of the patients, if dead, with vital status being complete to December 2014.

5.2 Derivation and interpretation of selected individual data items

5.2.1 Charlson comorbidity index

Derivation

Hospital Episodes Statistics (HES) data were used to generate the Charlson Comorbidity Index [103], for each patient at the point of diagnosis, based on Hospital Episode Statistics records from the period 6 to 60 months prior to their diagnosis using the algorithm created by Maringe *et al*, which was found to be a sufficient period to capture the vast majority of comorbidities recorded [104]. To calculate the index, a score is first assigned for each of a selection of pre-existing conditions the patient has, with higher scores given for more severe conditions (e.g. chronic lung disease = 1; AIDS = 6). The patient's Charlson score is then calculated as the sum of the scores for each of the conditions they have. I applied the algorithm developed by Maringe *et al* to the hospital records of the cancer patients studied in this thesis, and then linked the comorbidities information derived was to the cancer registrations dataset.

Interpretation

The conditions included and scores given to each condition in the Charlson Index are selected to ensure that the Index is highly correlated with a patient's hazard of death within the next year [103]. The Charlson Index is commonly used to help a physician assess a patient's current frailty when evaluating the risks and benefits of aggressive treatment for a new condition, such as a cancer diagnosis. As the Index measure used in this thesis is derived entirely from conditions recorded in Hospital Episodes Statistics, it only includes diagnoses in NHS inpatients, outpatients, or A&E settings. The exclusion of primary care data is unlikely to affect the accuracy of the scoring, as the treatment included in the score require diagnosis and treatment in either an inpatient or outpatient clinic. However, the completeness of data on diagnoses in outpatient clinics is poor in HES [105], and some small residual error is to be expected due to this limitation and due to the exclusion of conditions which were treated privately, or which were never diagnosed prior to cancer diagnosis. The Index measure used in this thesis is therefore likely to slightly underreport the prevalence and severity of comorbidities at time of cancer diagnosis.

5.2.2 Deprivation

Derivation

Patient's level of deprivation was estimated using the 2015 edition of the English Indices of Multiple Deprivation (IMD) [106], which are regularly produced by the Ministry of Housing, Communities & Local Government. For each Lower Super Output Area (LSOAs: small geographic areas containing ~1,500 people each on average [107]) seven indices of the local population are estimated: income; employment; health and disability; education, skills and training; barriers to housing and services; living environment; and crime. Individual patients can be assigned an overall deprivation score based on the addition of the scores for these different characteristics (also known as 'domains') of the LSOA

they live in. The overall and domain-specific deprivation scores data considered in this thesis were retrieved from the Office for National Statistics, and linked to individual patient records using their postcode at time at diagnosis by Adrian Turculet .

Interpretation

The measure of deprivation selected for this thesis was the income domain of the IMD. This has been shown to be highly correlated with health outcomes, and using it avoids the conceptual problem of "mathematical coupling" from analysing differences in health outcomes by deprivation when the measure of health deprivation used is itself partly based on health outcomes (through the health and disability domain) [108]. The income deprivation domain is calculated as the percentage of people in the LSOA who are recipients of state benefits (including income support, jobseeker's allowance, pension credit, working or tax child credit if the household income is below 60% of the median, and supported asylum seekers [106]). It can be considered a proxy for the percentage of people in the area who have very low incomes. The data used to calculate the scores are from 2008, and so are contemporary with outcomes for patients diagnosed during 2008-2013.

Deprivation scores can be included in analyses as the raw IMD score, or grouped into quintiles based on the overall distribution. The quintiles approach divides the LSOAs into five equal sized deprivation groups, from least to most deprived. This approach is useful for tabulations, stratified analyses, and matching individual patients based on deprivation (for example, to adjust for confounding from deprivation). Quintiles have historically been used in analyses of net survival in England, where patient mortality is compared to expected mortality amongst people with that deprivation quintile (and also age, sex, year) in patient's region of residence [109]. One disadvantage of using quintiles is that there are considerable differences in the percentage of people claiming benefits between the LSOAs within each quintile, particularly within the two most deprived groups, and information on these granular differences is lost [106]. Additionally, as they are ecological measures for the proportion of people with very low incomes, IMD scores do not discriminate between people with average and very high incomes. However, to facilitate easily interpretable tabulations and parsimony with variables, and consistency with previously reported studies in England, the quintiles of deprivation were chosen for use in this thesis.

5.2.3 Route to diagnosis

Derivation

Route to diagnosis in the health service was obtained by linkage to the routes to diagnosis dataset. This dataset categorises each patient with a cancer registration into one of 8 routes using the algorithm generated by Ellis-Brookes *et al* [54], which itself draws on data from cancer registrations, cancer waiting times, and screening programmes to determine the first activity within the NHS which eventually led to diagnosis. Based on records of activity in NHS datasets up to 6 months prior to cancer diagnosis, and the sequence in which they occurred, every cancer diagnosed is categorised into one of "screen-

detected", "two week wait" (urgent GP referral), "emergency presentation" (A&E attendance or emergency GP referral), "GP referral" (non-urgent non-emergency GP referral), "inpatient elective" (originating with an existing inpatient episode), "other outpatient", "death certificate only" (cancer was only recorded on death certificate), or "unknown" (no information in any NHS records).

Interpretation

The distribution of routes to diagnosis varies between cancers, but regardless of the cancer the emergency presentation route is associated with considerably lower survival than any of the others [54]. The algorithm is limited in having no data on activity of patients treated privately, who are categorised into the "unknown" or "death certificate only" categories.

Routes to diagnosis is specific to the healthcare setting, but not a particular clinical presentation. In particular, not all patients with the "emergency presentation" route necessarily had an acute medical problem prior to diagnosis: it's possible they first presented to A&E with mild symptoms. One study evaluating 35 common cancers from 2006 to 2015 reported that the proportion of all emergency presentations originating in A&E attendance increased in this period, while the proportion originating in emergency GP referral fell [110]. It was hypothesised that many of the latter were replaced with urgent referrals on the two week wait pathway instead, suggesting that the prognostic value of the emergency route might change according to changes in typical patient pathways through the health services.

5.2.4 Stage at diagnosis

Derivation

Information on stage at diagnosis was retrieved from cancer registration, CAS, and the audit datasets: if data on stage for a patient was present in any one of these it was used. The linkage and derivation of a final stage category was completed using the strategy and algorithm of Benitez Majano *et al*, and also executed by her for this analysis [111]. In cases of conflicting stage information between the datasets, stage was set to the value reported in the audit datasets. The TNM tumour stage fields were used for each of the three cancers, taking the values I, II, III, IV, or missing [112].

Interpretation

The TNM staging system categorises every tumour into one of:

- Stage I: disease localised to the organ of origin and of small size and extension
- Stage II: disease of greater size and extension contained within the organ of origin
- Stage III: disease which has spread to lymph nodes, is far advanced within the organ of origin, or which has spread to an adjacent organ
- Stage IV: disease has metastasised (spread to a distant organ)

The TNM stage system is a useful means of summarising disease spread as it allows comparison of stage between different cancer types using a common categorisation system. It was designed to

discriminate between patients with different survival prospects [112], and so is practically relevant to both patients and clinicians. “Early” stage is often defined as stages I or II, when the disease is localised and definitive curative treatment is typically possible.

Historically stage information was very poorly recorded in cancer registrations, though completeness improved dramatically from 2000 onwards [57]. The reasons for stage being incompletely recorded, and reasons for improvements in recent years, can be inferred by reference to the literature and analysis of patient’s routes to diagnosis, as outlined below.

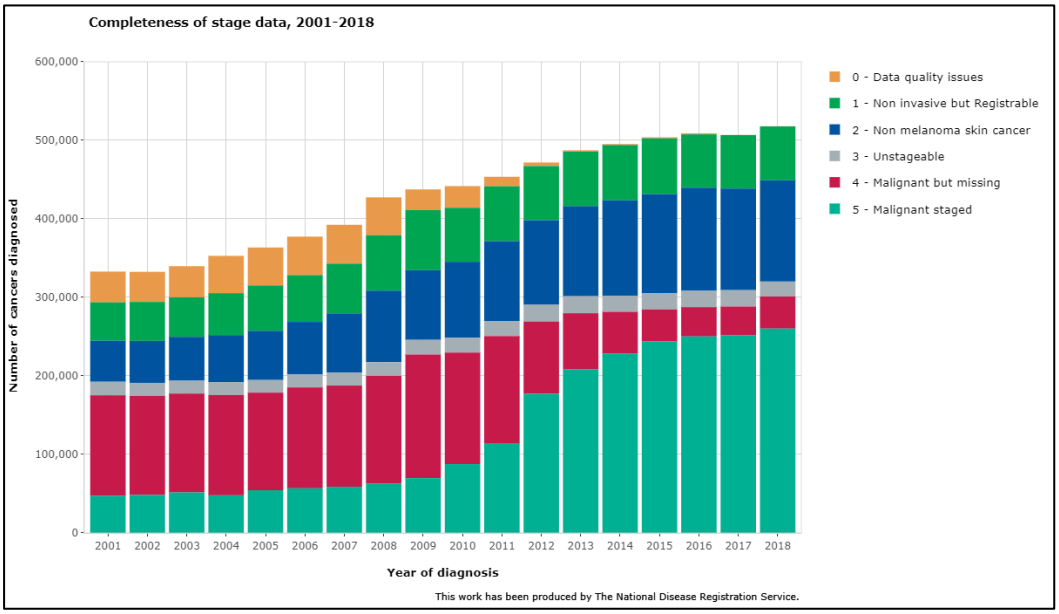
Public England categorise the staging status for each cancer registration into:

- *0 – The registration has data quality issues (ICD-10 code is either missing or unregistrable, or the registration is not finalised, or it is missing data on gender)*
- *1 – The tumour is non-invasive but registrable*
- *2 – Non melanoma skin cancer*
- 3 – Unstageable
- 4 – Malignant but missing
- 5 – Malignant staged

PHE define the *stageable* cancers (groups 4 and 5) as those for which a staging system exists for its combination of morphology and site (topography), and *unstageable* (group 3) as those for which a staging system does not exist for its morphology and site (topography) combination. In this thesis only registrations from the groups 3, 4, and 5 were extracted from the Cancer Registrations dataset (Figure 5.1).

Figure 5.1 shows that the composition of stage status of registered cancers changed dramatically during 2008-2013. The proportion of all stageable registered malignancies without stage recorded decreased, whilst the number categorised unstageable remained constant. For registrations with the ICD-10 codes for colon, rectum, lung, and ovarian cancers, an extremely small proportion of tumours are classed as unstageable (<150 records in 2013 for any of these cancers and typically much smaller numbers [57]), so the inclusion or exclusion of these tumours in the thesis analyses will have minimal impact on the results.

Figure 5.1 Distribution of staging data ascertainment categories for 21 common cancers in England, 2008-2013, from Public Health England's *CancerData* dashboard [57]

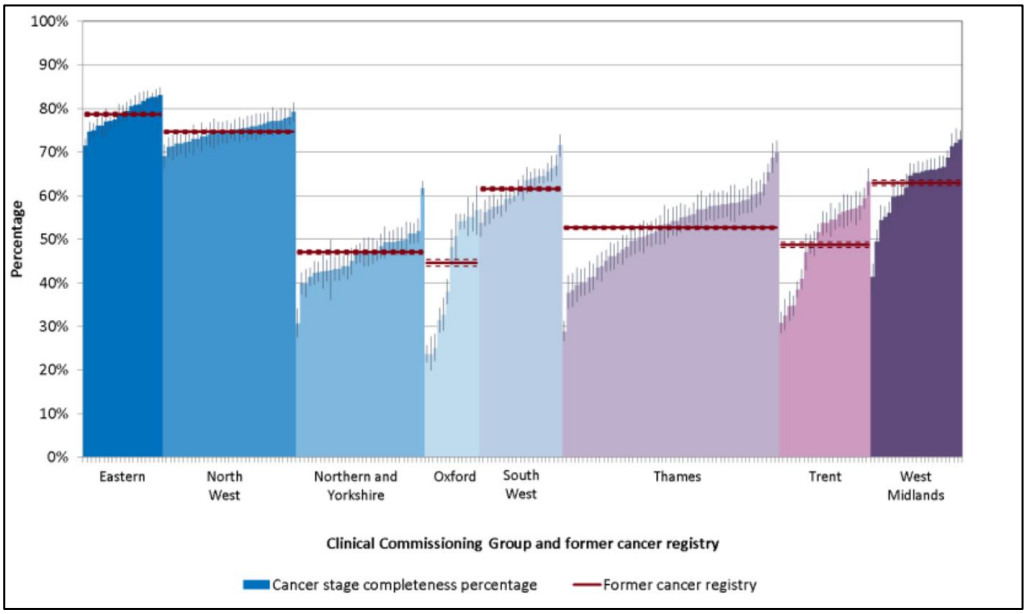


Consequently, where stage is missing in the cancer registrations analysed in this thesis, the tumour was of a type where stage could in principle have been recorded, but it was not. Four reasons why this occurs in England have been described by McPhail *et al* [113]:

1) It was not considered appropriate to stage the patient (as, for example, they would only ever be eligible for palliative care so ascertaining stage would not change the treatment plan);
2) The patient was diagnosed and treated outside the National Health Service;
3) The patient died before staging was complete;
4) Staging information was not transferred to or collated by the Cancer Registration Service.

Until 2013, when the National Cancer Registration Service took responsibility for Cancer Registration, it was conducted by 8 regional cancer registries. Each of these had different practices and each led on the registration of different malignancies. There are thus large variations in percentage of cancers staged between the registry areas up to 2013 (differences in 2012 in Figure 5.2, taken from the National Cancer Intelligence Network [114]). These differences along regional registry lines suggest a large administrative component to geographic differences in staging completeness. This is supported by one analysis of patients missing stage data in 2013, which reported that patient characteristics alone cannot explain the between-CCG variation in stage recording [115].

Figure 5.2 Completeness of staging data for all cancers by CCG and by historic regional Cancer Registry area in 2012, taken from the National Cancer Intelligence Network [114].



Reports from the Department of Health and Cancer Registries indicate that the increase in stage recording during 2008-2013 was driven by administrative and technical improvements in data transfer and collation, rather than underlying changes in the number of cancers for which staging information was ascertained in the hospital. The 2007 *Cancer Reform Strategy* states stage is only variably recorded by the MDTs treating patients, and that they must be obliged to record this information and send it on to the Cancer Registries [2]. Annual reports from the United Kingdom and Ireland Association of Cancer Registries (UKIACR) locate the reason for improvements in stage completeness in administrative practices and resourcing. For example, the commentary for England from the 2014 report (for registrations from the year 2012) notes:

"The most significant change from last year has been the dramatic improvements in staging rates for invasive cancers: moving from 51% last year to 62%. Within this, the NCRS has recorded stage rates of above 80% for the priority cancer sites of breast, colorectal, lung, ovarian, and prostate. It has done so by acting as a cohesive national service working on a single system." [116]

Similarly, the UKIACR report for 2015 discusses the variability in completeness of stage recording between regions, and recent improvements in stage completeness in Northern Ireland resulting from increased investment:

"Staging data has improved remarkably but there is still more work to be done. In areas where countries have focused their resources, the PI [performance indicator] results are strong (for example, Northern Ireland's staging completeness performance). NICR [Northern Ireland Cancer Registry] were able to achieve increases in staging data this year as a result of additional staffing resources. This highlights a need in all registries to consider the resources required to enhance staging completeness in future." [117]

Therefore, for an unknown but presumed large proportion of the cancer registrations where stage is not recorded, administrative factors are the reason (i.e. the data was not input and delivered to the registry): stage was still ascertained to guide a treatment decision in the hospital. Additionally, in these documents the improvements in recording from 2008 are attributed to administrative improvements in stage data collection, not increases in stage ascertainment at the hospital. These details can help analysts make more reasonable assumptions about the likely true distribution of stage for the cases where it is not recorded in the registration.

Comparison of the routes to diagnosis of patients missing stage information can also inform assumptions about the distribution of stage where it is not recorded. A tabulation of route to diagnosis for patients with and without stage recorded is presented in Table 5.1.

Table 5.1 Distribution of patients by year and route to diagnosis, for all patients and patients whose stage was not recorded or missing in the cancer registration, England, 2008-2013.

	All patients with a cancer registration (with or without stage recorded)		Patients whose stage was not recorded in the cancer registration	
	2008	2013	2008	2013
Colorectal cancer (n=196,516)				
Emergency presentation, %	23.2	23.3	23.6	29.5
GP referral, %	23.6	22.5	23.3	20.0
Inpatient elective, %	4.9	3.3	4.6	2.8
Other outpatient, %	8.0	6.6	8.2	6.5
Screening , %	4.8	9.5	4.9	5.8
Two week wait, %	26.4	30.1	23.7	16.9
Unknown, %	9.2	4.6	11.9	18.5
All routes, %	100.0	100.0	100.0	100.0
NSCLC (n=180,230)				
Emergency presentation, %	36.8	34.1	44.9	52.0
GP referral, %	21.3	21.4	20.1	17.6
Inpatient elective, %	1.7	1.6	1.8	1.6
Other outpatient, %	10.4	11.7	10.4	9.6
Two week wait, %	24.6	27.9	15.4	6.8
Unknown, %	5.1	3.4	7.5	12.4
All routes, %	100.0	100.0	100.0	100.0
Ovarian cancer (n=29,221)				
Emergency presentation, %	32.5	29.5	36.2	42.7
GP referral, %	22.2	21.5	21.7	20.5
Inpatient elective, %	1.8	0.9	1.9	1.0
Other outpatient, %	12.4	9.6	12.6	10.1
Two week wait, %	24.1	35.7	20.1	20.8
Unknown, %	6.9	2.8	7.6	4.9
All routes, %	100.0	100.0	100.0	100.0

Three key patterns stand out in this table. First, patients whose stage was not recorded were more likely to be diagnosed following emergency presentation than patients whose stage was recorded, and this disparity was greater in 2013 than in 2008. For example, in 2013 52% of NSCLC patients whose stage was not recorded were diagnosed following emergency presentation, compared to only 34% of all NSCLC patients. Second, patients whose stage was not recorded were also more likely to have an unknown route to diagnosis than patients whose stage was recorded. This disparity was also greater in 2013 than in 2008 for colorectal cancer and NSCLC. Third, despite these differences, patients with and without stage recorded had broadly similar routes to diagnosis, though their routes were less similar in 2013 than in 2008.

Considering the previously discussed four possible reasons why stage may not be recorded for a patient (not appropriate; stage recorded outside the NHS; the patient died; not collated or transferred to the Cancer Registry), these results indicate each reason contributed to stage being missing in 2008-2013. That the routes to diagnosis of patients with and without stage recorded are generally similar is consistent with stage typically being missing for administrative, rather than clinical, reasons. However, the emergency and unknown routes were more common for patients without stage recorded, and this is to be expected if stage is occasionally not recorded either because it was not ascertained; or because the patient died; or they were treated privately. Finally, given the administrative improvements in stage recording between 2008 and 2013, it is expected that the other three reasons should account for a proportionately larger share of the total by 2013. This is also suggested by the larger proportion of diagnoses via unknown or emergency routes amongst patients without stage recorded in 2013 compared to 2008.

In summary it is likely that stage at diagnosis was historically primarily missing in cancer registration records for administrative reasons, unrelated to patient's outcomes. However, it also likely that in the periods after administrative improvements in stage data collection by the Registries had occurred, the patients without stage recorded were increasingly those for whom stage was never ascertained within the NHS because they were frail; because they died prior to staging being complete; or because they were treated privately. In the context of minimising bias from missing stage, it being missing purely for administrative reasons unrelated to patient care or outcomes implies a "missing completely at random" mechanism: systematic differences in average stage between patients with stage recorded or stage not recorded are not expected. However, in practice stage can be missing for the other reasons, and in those cases the patients with later stage disease may be less likely to have it be recorded. This implies that stage may only be missing randomly conditional on other factors, or not missing at random.

5.2.5 Tumour morphology

Derivation

Tumour morphology coded according to the International Classification of Diseases for Oncology (3rd edition) was retrieved from the Cancer Registration datasets [118].

Interpretation

Morphology includes the histological characteristics of the tumour (the structure of the cells and tissue), whether the tumour behaviour is malignant (of a type that will continue spread) or benign (not of a type that will spread); and, for malignant tumours, the tumour grade (how fast growing it is). It is ascertained by pathological examination of a sample of the cancerous tissue. Morphology coded following the ICD-O standards contains the following pieces of information [118]:

- **Cell type (4 digit code):** The cell type the tumour originated in.
- **Behaviour (1 digit code):** 0=benign; 1=uncertain/borderline; 2=in situ, non-invasive; 3=malignant, primary site; 6=malignant, metastatic/secondary site; 9=malignant, uncertain whether primary or metastatic.
- **Grade (1 digit):** How closely the tumour cells resemble normal cells from that tissue; well differentiated (low grade) cells resemble normal cells and are slow-growing, whilst poorly-differentiated cells are very abnormal and faster-growing (high grade). Grade is coded as follows: X=cannot be assessed, 1=well differentiated (low grade), 2=moderately differentiated (intermediate grade), 3=poorly differentiated (high grade), 4=undifferentiated (high grade).

Table 5.2 Distribution of recorded stage by recorded grade in cancer registrations in England, 2008-2013.

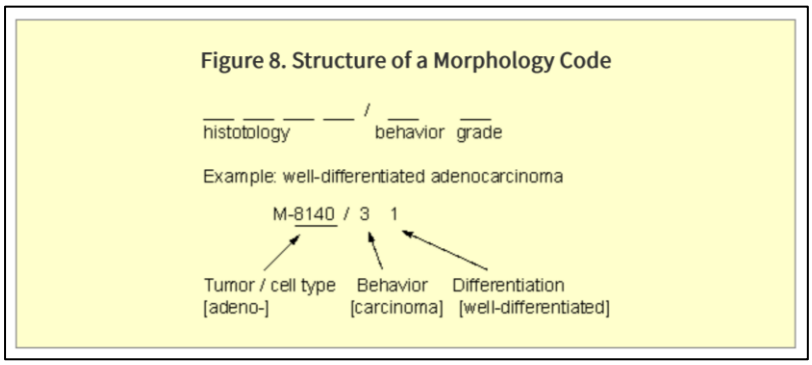
	Grade 1, %	Grade 2, %	Grade 3, %	Grade 4, %	Missing, %	Total, %
Colorectal cancer (n= 196,516)						
Stage 1	9.6	72.6	6.1	0.1	11.7	100.0
Stage 2	4.3	75.6	11.3	0.1	8.6	100.0
Stage 3	3.3	67.4	17.8	0.2	11.3	100.0
Stage 4	2.8	46.2	16.1	0.3	34.5	100.0
Missing	4.9	50.9	10.6	0.1	33.6	100.0
Total	4.6	58.5	12.5	0.2	24.2	100.0
NSCLC (n=180,230)						
Stage 1	5.5	20.0	16.1	0.5	57.8	100.0
Stage 2	3.1	19.4	23.2	0.9	53.5	100.0
Stage 3	1.6	12.5	21.0	0.8	64.1	100.0
Stage 4	1.2	7.5	18.4	0.9	72.0	100.0
Missing	1.4	5.3	9.5	0.7	83.0	100.0
Total	2.0	10.3	17.2	0.8	69.8	100.0
Ovarian cancer (n=29,221)						
Stage 1	20.9	18.8	20.3	0.4	39.5	100.0
Stage 2	5.7	10.8	39.3	0.7	43.5	100.0
Stage 3	2.5	5.8	39.8	1.0	50.9	100.0
Stage 4	1.5	5.9	33.2	1.0	58.5	100.0
Missing	3.8	6.5	21.3	0.5	67.9	100.0
Total	5.9	8.3	28.0	0.7	57.0	100.0

Only invasive malignancies (behaviour=3) were selected from the cancer registrations, since benign/uncertain/in-situ tumours are not definitely hazardous to health or within the scope of early diagnosis interventions.

Information on grade was generally poorly and variably recorded in cancer registrations in England during 2008-2013. Grade was missing for 24.4% of colorectal cancer patients overall and 33.7% of patients without stage recorded, compared to 70.6% and 83.7% of NSCLC patients respectively, and 57.7% and 68.4% of ovarian cancer patients (Table 5.2).

Information on tumour cell type is well recorded by contrast, and this is also associated with tumour grade. For example, mucinous, endometroid and clear cell ovarian carcinomas are slow growing compared to other histologic types [119]. In England in 2008-2013, 11% of colorectal cancer tumours were registered with a “non-specific” or “unknown” morphology cell type, compared to 10% of NSCLC tumours and 6% of Ovarian cancer tumours (further discussion in Chapter 10).

Figure 5.3 Typical structure of morphology codes reported in line with the ICD-O from the USA Surveillance Epidemiology and End Results (SEER) programme [120]



A substantial minority of the patients in the English datasets are coded as having ‘non-specific’ or ‘misc. and un-specified’ morphologies. These are codes which are assigned when the morphology cell type has not been ascertained. Ascertainment requires the patient to have undergone surgery or a biopsy to extract a sample of the cancerous tissue, and for microscopic verification to have been performed by a pathologist. Patients who are acutely unwell at the time of diagnosis, or who have poor prognosis, may be less likely to undergo these procedures (due to lack of time before death, physical frailty, or futility) and hence categorised as having ‘non-specific’/’unspecified’ morphologies, and also lack information on grade in the morphology code.

5.2.6 Tumour topography

Derivation

Tumour topography describes the anatomical organ and sub-tissue the tumour originated in. Like the morphology data, it is categorised using the ICD-O classification, which itself takes topography codes

from the generic international standard ICD-10, and was retrieved form the Cancer Registrations supplied by the ONS [121].

Interpretation

Topographical sub-site is associated with stage at diagnosis for each of the cancers considered in this thesis. For example, right colon tumours typically initially present with non-specific symptoms compared to rectal tumours, and cannot be detected by sigmoidoscopy [65], and hence are harder to diagnose early.

A substantial minority of patients have the tumour coded with 'Other'/'Not otherwise specified (NOS)' topography (full details for each cancer analysed are presented in Chapter 10.2). Patients presenting with late-stage disease or a medical emergency may be less likely to undergo a full diagnostic investigation, and, for those with late stage disease, it may be more difficult to determine the exact topographical sub-site the tumour originated in, as it is no longer localised. Therefore, similarly to morphology, the non-specific codes may arise from late stage at diagnosis.

5.2.7 Major surgery following diagnosis

Derivation

Hospital Episodes Statistics records were used to provide information on receipt of major surgical treatment for the patient's cancer following diagnosis. In brief, for each of the three tumour sites, OPCS Classification for Interventions and Procedures version 4 codes for surgical interventions were separately extracted and used to identify operations which would have been performed in order to remove cancerous tissues in admitted patient care during from one month prior to diagnosis to six months after it (full list of procedures is in Chapter 11, Appendix Table 6 [122, 123].

Interpretation

These records have one limitation similar to the routes to diagnosis data: they exclude patients treated privately. Otherwise, they are expected to be complete, barring coding errors in the hospital, since they are based on the same records used to compensate NHS Trusts based on volumes of activity recorded under the *Payment by Results* system [105].

5.3 Summary

Numerous population-based health data sources were linked to cancer registrations to provide extra information on patients diagnosed with colorectal, NSCLC, or ovarian cancer during 2008-2013. Together these data provide a wealth of information on the tumour spread and behaviour, on the demographics of the patients and their activity within the health system around the time of diagnosis, and patient's vital status from date of diagnosis to 31 December 2014.

Some important considerations and limitations for certain data items were identified. Data on pre-existing comorbidities and major surgery were derived from NHS secondary care, and so miss activity from primary care or the private sector. The deprivation measure is based on characteristics of patient's area of residence, and so is only a proxy for their individual circumstances.

A big challenge for monitoring of early diagnosis is that stage is not recorded in the registrations of a large proportion of cancer patients during 2008-2013. Evidence from the literature and an initial data analysis for the three cancers evaluated in this thesis indicate stage is primarily missing for administrative reasons (unrelated to patient's clinical situation), but occasionally missing because the patient was too frail for staging to be in their interest; because they died; or because they were treated privately. Therefore, the patients with and without stage recorded can't be assumed to have exactly the same stage distribution, particularly in 2013 after administrative improvements in data collection.

There is a similar challenge for data on topographical subsite and morphology. These may be coded as missing or unspecified, and these categories are *a priori* more likely in patients who are acutely unwell (in some cases, due to advanced disease). In the context of case-mix adjustment, these missing or unspecified values may be a proxy for late-stage disease or another acute problem. Hence, rather than providing information on case mix, they could be a consequence of late diagnosis, and adjustment for them may mask early diagnosis differences. As with stage, an analytical strategy for handling these missing data is needed to minimise bias in the evaluation of trends.

6. A conceptual framework for early cancer diagnosis

One key aim of this thesis is to produce statistics to evaluate health service performance in early diagnosis during 2008-2013. These statistics should ideally take account of those patient and tumour factors which affect the likelihood of early diagnosis, but which are not within the control of health services. In this Chapter, these factors are identified and described with reference to the literature, and placed within a conceptual framework. This conceptual framework is then related to the data items actually available for the analyses, and used to identify variables to be adjusted for in the substantive analyses which follow.

The frameworks in this Chapter are similar to those used in epidemiological causal inference studies [124]. However, construction of a causal framework to minimise confounding in an estimate of the causal effect of one exposure on an outcome is not the objective here. Health service effectiveness, the key exposure, is never measured directly, but is assumed to be the dominant factor determining early diagnosis trends after other factors which affect propensity for it have been accounted for. The rationale for identifying factors for adjustment is thus the same here as for national cancer survival estimates and other official health statistics; to make comparisons more meaningful by "*taking account of patient characteristics and factors which are beyond the control of providers*" [125]. Comparison of adjusted and un-adjusted results may also allow some inferences to be made about the role of case mix in explaining observed trends.

6.1 Defining early cancer diagnosis

"Early diagnosis" typically refers to "early-stage disease at time of diagnosis". However, as it has become possible to link the electronic health records of cancer patients to their cancer registrations, researchers are increasingly using other indicators to measure how prompt diagnosis is (or if there have been failures in diagnosing patients early) based on these health records instead of stage. Some of these indicators include the interval from first symptoms to diagnosis, emergency presentation, and (all-cause) mortality shortly after diagnosis. These other indicators can give insight into the circumstances of patients whose stage was never ascertained or reported to the cancer registry.

However they are defined, early diagnosis statistics have the potential advantage of being more rapidly available than survival statistics. The one-year survival statistic presented on the *CancerData* dashboard in 2020 are for patients diagnosed in 2017, reflecting the time needed to achieve a year follow-up, then conduct and publish the survival analyses [5]. By contrast early diagnosis statistics for 2019 are presented: two years more contemporary. Early diagnosis statistics could, in principle, be published in a matter of weeks or months after the reporting period ends, allowing rapid feedback on the effect of different interventions or events.

There has been debate amongst cancer epidemiology researchers in England regarding which indicators should be included in surveillance, and how the statistics derived from them should be specified [44]. Some topics in that debate include which indicators are most informative and highly associated with survival, and how missing data should be handled. For the purposes of the framework for case-mix adjustment presented in this Chapter, emergency presentation and mortality are included as well as stage. Interval from first symptoms to diagnosis is not included, as the indicator has a complex relationship with other indicators and survival (discussed further in Chapter 7).

6.1 A conceptual framework for determinants of early cancer diagnosis

Determinants of early diagnosis can be divided into two groups conceptually. The first group are factors of health service performance:

- Screening effectiveness and uptake;
- Patient education and awareness;
- Patient's propensity to attend primary care with symptoms;
- Responsiveness of the primary care services to patients;
- Promptness of pathways from primary to secondary care;
- Quality and timeliness of diagnostic tests.

The second group are tumour and patient factors:

- Tumour site and sub-site of origin
- Tumour morphology (including grade)
- Patient's age and sex
- Comorbidities
- Deprivation (whether income, social, or educational)

These two groups of factors interact with each other to determine the length of time from tumour development to formal diagnosis, which in turn determines the extent of disease spread at diagnosis, and whether or not the patient is diagnosed due to a medical emergency. The relationships between the different patient factors and early diagnosis are as follows:

Tumour site and subsite

The probability of early diagnosis varies markedly depending on the tumour's anatomical location of origin and subsite within the organ [48, 97, 126, 127], partly because the stage a cancer reaches before it produces specific symptoms varies by location. The anatomical site mix of newly incident cancers in an area reflects both current demographic and historical characteristics in the local population. For example, the proportion of incident lung cancers within an area in a given year is largely determined by the prevalence of smoking in its population in previous decades [128]. For

comparisons of health services performance, the distributions of tumour site (and subsite) are therefore best viewed as confounding factors to be controlled for.

The probability that a patient requires urgent medical treatment prior to diagnosis also depends on where the cancer arises, as different tumour sites have varied associated complications. For example, right-side colon or advanced ovarian cancer may result in a bowel obstruction prior to diagnosis [72], while lung cancer can result in the partial blockage of major airways or severe bleeding. Tumour subsite is therefore also associated with emergency presentation and short-term mortality risk.

Tumour morphology

Higher grade has been shown to be associated with late-stage disease for colorectal cancer [129]. Tumour grade differences have also been posited as explaining the 'waiting times paradox', in which the patients with the shortest waiting times have the poorest outcomes. It is hypothesised that the aggressiveness of their disease causes some patients to present as acutely unwell or with red-flag symptoms, which leads both to a quicker diagnoses and poorer outcomes compared to patients with lower-grade disease [130].

Age and sex

Patient's age and sex are strongly associated with the probability of early diagnosis [56, 126, 131]. Increasing age is associated with late diagnosis in part because the increasing frailty and co-morbidities of older patients make it less likely that cancer symptoms will be promptly identified and investigated. Some sex differences in late diagnosis may be due to false attribution of symptoms from cancer to other conditions which are sex-specific or more prevalent in one sex. For example, abdominal pain is more common amongst women and also a symptom of colorectal cancer, and women are more likely to be diagnosed followed by emergency presentation than men for that cancer [126]. Social factors, such as discrimination based on age and sex, or indirect barriers in accessing health services associated with age or sex [132], may also play a large role in explaining these differences.

Pre-existing comorbid conditions

Co-morbidities have a complex relationship with early diagnosis, and may increase or decrease it depending on the tumour site and comorbidity. Certain comorbidities are associated with later diagnosis, potentially by creating barriers to accessing health services or misattribution of cancer symptoms to other causes [133]. There is also evidence that specific co-morbidities may aid earlier diagnosis in some cases. For example, a large proportion of early-stage lung cancers are diagnosed incidentally, when the patient is undergoing a chest X-ray for another condition [134], and so pre-existing cardiovascular disease (which requires an X-ray) is associated with earlier-stage lung cancer diagnosis [135].

Deprivation

Patient deprivation is associated with both late-stage diagnosis, emergency presentation, and lower survival for several common cancers [126, 136, 137]. However, for lung cancer a link between higher deprivation and later-stage disease has not been found in England [138, 139]. Deprivation can be linked to late diagnosis through lower health-seeking behaviour and co-morbidities, but also through health services factors: the health system may be more difficult to navigate for deprived patients [140], and services located in areas of high deprivation may have additional financial and operational pressures as a result of their patient case mix or problems recruiting key staff.

Summary

Tumour site, subsite, and grade, patient's age and sex, and comorbidities, are associated with probability of early diagnosis, but are generally unrelated to the performance of health services at a given point in time. These factors should be adjusted for or otherwise controlled for when comparing the performance of health services in early diagnosis. Deprivation may additionally be adjusted for to account for the increased challenges to health services in areas of high deprivation. However, deprivation-adjusted statistics may give a misleading impression of the extent of *avoidable* inequalities, so care must be taken when interpreting these.

Figure 6.1 shows an *a priori* conceptual framework for the interactions between these different factors and early diagnosis outcomes. In the framework, patient's frailty and co-morbidity are influenced by both age and sex, by patient's deprivation [141], and by education and health seeking behaviour [142]. The probability of early diagnosis is hypothesised to be higher for patients who are better educated and who exhibit health seeking behaviour, and to vary for patients who are frail or who have co-morbidities (either lower or higher depending on the specific cancer and comorbidity). The risk of an acute medical problem prior to diagnosis is higher with (undiagnosed) later stage disease, and also if the patient is frail or has comorbidities [143]. Finally, patient's risk of death shortly after diagnosis is determined by the disease stage at diagnosis, but also by their frailty and comorbidities [103], whether the patient had an acute medical problem prior to diagnosis [96], and by the performance of the health services at the point of presentation. Health system effectiveness is also assumed to have direct influence over a patient's level of health education and health seeking behaviour, and an effect on the disease stage at diagnosis.

Figure 6.1 *A priori* underlying conceptual framework for relationships between patient and health services attributes (yellow), and selected outcomes (blue) for early cancer diagnosis, for a tumour originating in a given anatomical site.

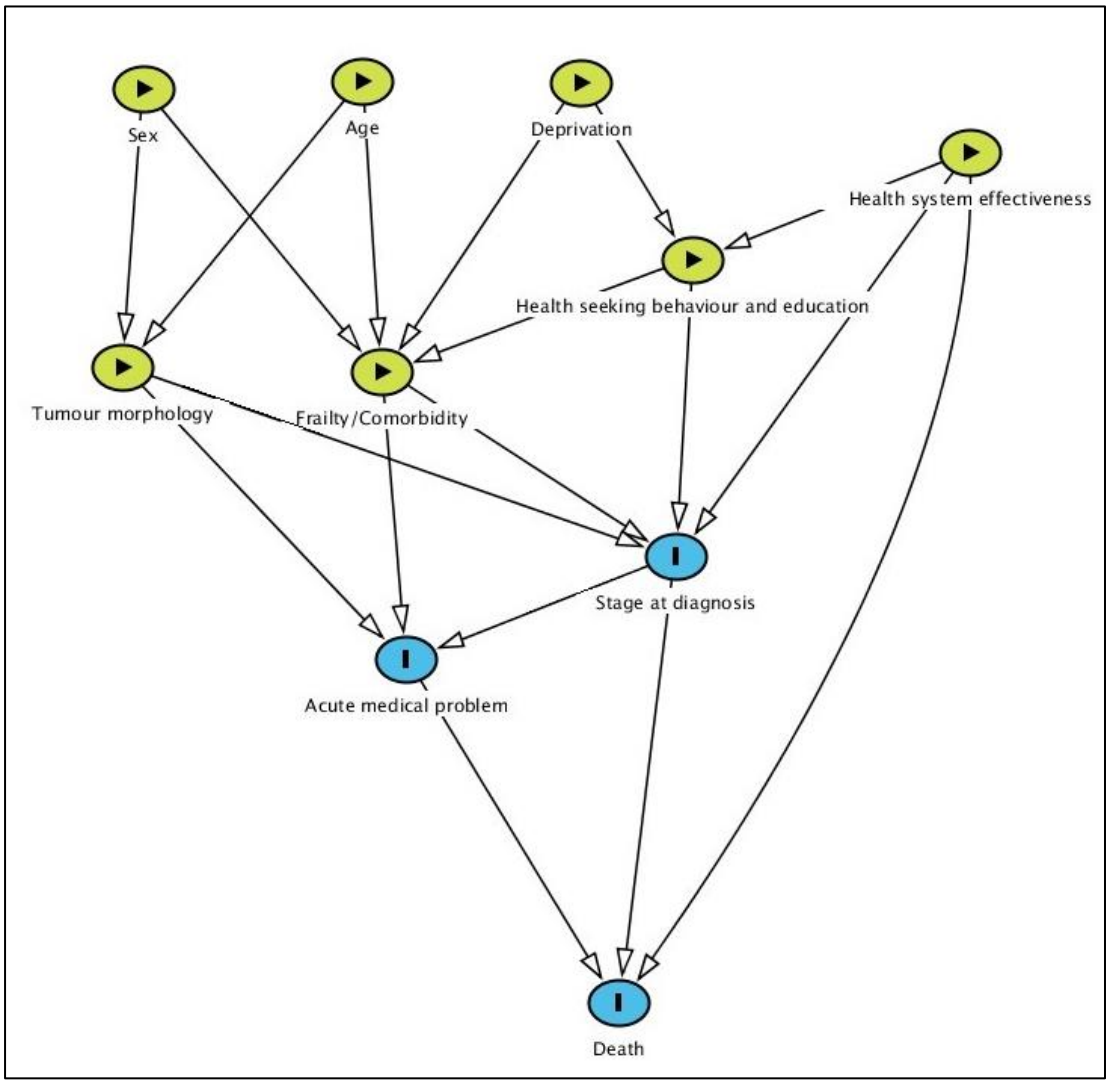
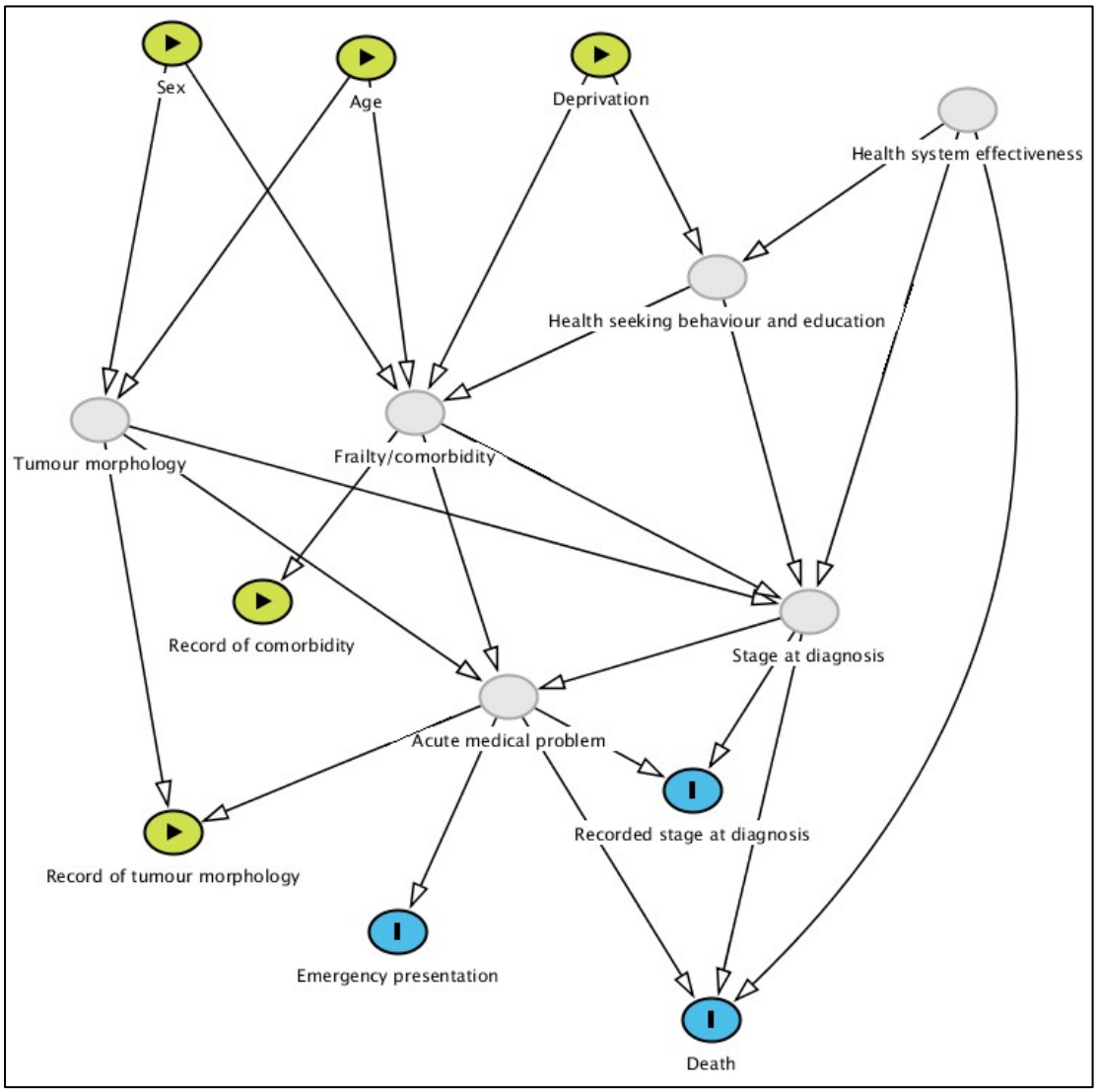


Figure 6.2 Conceptual framework for relationship between recorded exposures (yellow), latent un-recorded exposures (grey), and selected recorded outcomes (blue) for late cancer diagnosis, for a tumour originating in given anatomical site.



6.2 A framework with population-based data items

As discussed in Chapter 5, data for some of the factors in the framework presented in Figure 1 are either not recorded or incompletely recorded in the data sources used in this thesis:

- Only co-morbidities which were diagnosed or treated in NHS secondary care services, and which were correctly recorded there, are available for analysis;
- Stage at diagnosis was not recorded for >10% of patients with common cancers diagnosed in 2013, and much higher proportions in previous years [5]. Some of these patients may lack a

recorded stage because they had an acute medical problem or had poor prognosis at the point of presentation, potentially as a result of late-stage disease;

- Patients who had an acute medical problem prior to diagnosis are very likely to have the 'emergency presentation' indicator, however it is not specific only to the patients who experienced an acute medical problem prior to diagnosis;
- Tumour morphology is poorly recorded, being missing for 24%, 70%, and 57% of all patients with colorectal, NSCLC, and ovarian cancer respectively during 2008-2013 (as described in Table 5.2).
- Information on individual health education level, and health seeking behaviour, are not available in the population-based healthcare data sources used in this thesis. The average income deprivation in a patient's area of residence is likely to be only a very rough proxy measure for these.

Figure 6.2 shows the pathways between the variables which are partially available in population-based routine datasets (yellow), latent un-recorded variables (white), and recorded early diagnosis outcomes (blue). Recorded stage at diagnosis can take the values: I, II, III, IV, or missing, with value taken depending both on actual stage at diagnosis, and by how unwell the patient was at point of diagnosis. Similarly, recording of morphology and topographical subsite also depend on the patient's condition at time of diagnosis.

6.3 Implications for case-mix adjustment and the selection of an early diagnosis indicator for monitoring

The conceptual frameworks in Figures 6.1 and 6.2 help identify factors to include in case-mix adjustment for the analyses in this thesis: only those which predict early diagnosis independently of any effect from health services performance. Variables downstream from the early diagnosis indicator used should also not be adjusted for, as these are not case-mix factors and adjustment would mask the differences which are of interest. For example, mortality may be a consequence of late stage disease, and so should not be adjusted for in an analysis of stage trends.

Based on the first framework, age, sex, tumour morphology, and comorbidities should ideally be included as case-mix variables. Though the prevalence of comorbidities may be partly determined health system effectiveness, and so could arguably not be adjusted for as a case-mix factor, many of the most common comorbid conditions take years or decades to develop (e.g. heart disease, emphysema) so the prevalence of these in a given year is not relevant to evaluating cancer service effectiveness in that year.

Figure 6.2 highlights factors which can't be adjusted for because the data is not available. Patient's health educational level, and their health seeking behaviour, are not measured in population-based data sources. Deprivation is only a weak proxy for these.

It is also debatable whether deprivation itself should be adjusted for. Adjustment hides inequalities in health outcomes associated with deprivation, which can imply these are “warranted” or expected. However, many of the aspects of deprivation which affect early diagnosis (e.g. lower health-seeking behaviour; staffing problems and stretched resources; social obstacles to engaging with health services staff) are in principle modifiable through actions of the health services. On the other hand, practically, the resources needed to effect such changes may not be available. For the thesis analyses deprivation was not adjusted for, so that any geographic inequalities associated with deprivation would not be masked. As deprivation is not adjusted for, as with patient health-seeking behaviour and education level, it needs to be considered as a potential explanatory factor for differences observed.

Figure 6.2 illustrates how failure to record stage and morphology may actually be a consequence of late diagnosis. As discussed briefly in Chapter 5, these missing values must be handled appropriately: simply defining “missing” as another value the categorical case-mix variable can take would result in adjustment for something which is downstream of early diagnosis, masking differences in it. Figure 2 also illustrates some of the challenges associated with use of different early diagnosis indicators in monitoring: ‘Emergency presentation’ is not a specific indicator of acute medical problems, and stage at diagnosis is not completely recorded. The interpretation of the different early diagnosis indicators, and the association between each early diagnosis indicator and survival, is evaluated in Chapter 7, leading to further discussion about the most appropriate to use in the substantive analysis.

7. Which indicators of early cancer diagnosis from population-based data sources are associated with short-term mortality and survival?

In order to meet the first thesis aim of identifying the most informative indicator for monitoring early diagnosis trends, indicators were surveyed through a systematic literature review, and evidence for their association with patient survival was extracted. An empirical data analysis measuring the association between each indicator and patient's short- and long-term survival was also conducted. The combined literature review and data analysis was published in the journal *Cancer Epidemiology* in 2018 and is presented here in full. This Chapter concludes with further discussion of the results, leading to the decision and justification of selecting stage as the indicator for the substantive analysis.

7.1 Published manuscript

7.2 Supplementary appendices

Appendix Table 1 Literature search keywords and document inclusion criteria.

Elements*	PubMed search terms**	Google search terms***
cancer	cancer* OR tumor* OR tumour* OR malignan* OR neoplasm*	cancer OR tumor OR tumour OR malignancy OR malignancies OR neoplasm
early diagnosis	"early cancer diagnosis" OR "earlier cancer diagnosis" OR "early diagnosis" OR "earlier diagnosis"	early cancer diagnosis OR earlier cancer diagnosis OR early diagnosis OR earlier diagnosis
population-based	"population-based" OR "population based" OR routine data* OR "routinely collected"	population-based OR population based OR routine data OR routinely collected
Document inclusion criteria		
Criterion (1)	Includes presentation of new statistics or methods for generating statistics based on an explicitly defined indicator of early diagnosis from data sources that are or could be collected routinely	
Criterion (2)	Includes results describing the association between an early diagnosis indicator and patient prognosis (any measure of complications, mortality, survival, life expectancy, or cure).	
* "AND" operator was applied to only return documents containing all three elements **Title/abstract search ***Different syntax was required to account for the differences in search options and logic implemented between PubMed and Google searches		

Appendix Table 2 Number of patients (%) diagnosed with colorectal, non-small cell lung, or ovarian cancer by route to diagnosis, England, 2009-2013.

Route to diagnosis	Colorectal cancer	Non-small cell lung cancer	Ovarian cancer
Death Certificate Only (DCO)	70 (0.0)	163 (0.1)	15 (0.1)
Emergency presentation	37,285 (23.2)	60,632 (35.6)	7,503 (30.7)
GP referral	38,253 (23.8)	35,357 (20.7)	5,291 (21.6)
Inpatient elective*	5,800 (3.6)	2,762 (1.6)	282 (1.2)
Other outpatient†	11,488 (7.2)	18,338 (10.8)	2,597 (10.6)
Screening	15,046 (9.4)	-	-
Two-week wait (TWW)	45,898 (28.6)	45,809 (26.9)	7,485 (30.6)
Unknown‡	3,939 (2.5)	3,352 (2.0)	734 (3.0)
Missing	2,838 (1.8)	4,022 (2.4)	543 (2.2)
Total	160,617 (100.0)	170,435 (100.0)	24,450 (100.0)
* A route to diagnosis commencing with a planned inpatient admission † An elective route starting with an outpatient appointment ‡ No records of inpatient, outpatient, or screening activity prior to diagnosis			

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	270040	**Title**	Mr
First Name(s)	Patrick		
Surname/Family Name	Muller		
Thesis Title	Statistical approaches for monitoring early cancer diagnosis in England		
Primary Supervisor	Dr Laura Woods		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Cancer Epidemiology		
When was the work published?	July 2018		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I was the lead author of this study. I planned the study, conducted the analysis, and prepared the draft of the paper. Co-authors provided feedback and input on the study design and the manuscript draft.

SECTION E

Student Signature	P Muller
Date	10/05/2021

Supervisor Signature	L Woods
Date	10/05/2021

Cancer Epidemiology 56 (2018) 161–170

Contents lists available at ScienceDirect

Cancer Epidemiology

journal homepage: www.elsevier.com/locate/canep

Which indicators of early cancer diagnosis from population-based data sources are associated with short-term mortality and survival?

Patrick Muller*, Sarah Walters, Michel P. Coleman, Laura Woods

Cancer Survival Group, Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

ARTICLE INFO

Keywords:
Cancer mortality
Surveillance
Cancer survival
Diagnostic interval
Early detection
Early diagnosis
Emergency presentation
Population-based
Routine data
Stage at diagnosis

ABSTRACT

Background: A key component of recent English cancer policy is the monitoring of trends in early diagnosis of cancer. Early diagnosis can be defined by the disease stage at diagnosis or by other indicators derived from electronic health records. We evaluate the association between different early diagnosis indicators and survival, and discuss the implementation of the indicators in surveillance of early diagnosis.

Methods: We searched the PubMed database and grey literature to identify early diagnosis indicators and evaluate their association with survival. We analysed cancer registrations for 355,502 cancer patients diagnosed in England during the period 2009–2013, and quantified the association between each early diagnosis indicator and 30-day mortality and five-year net survival.

Results: Each incremental difference in stage (I–IV) predicts lower 5-year survival, so prognostic information is lost in comparisons which use binary stage indicators. Patients without a recorded stage have high risk of death shortly following diagnosis and lower 5-year survival. Emergency presentation is independently associated with lower five-year survival. Shorter intervals between first symptoms and diagnosis are not consistently associated with improved survival, potentially due to confounding from tumour characteristics.

Interpretation: Contrary to current practice, we recommend that all the stage information should be used in surveillance. Patients missing stage should also be included to minimise bias. Combined data on stage and emergency presentation could be used to create summary prognostic measures. More work is needed to create statistics based on the diagnostic interval that will be useful for surveillance.

1. Introduction

Increasing early-stage diagnosis is a common component of regional and national strategies to reduce the burden of cancer [1–5]. 'Early diagnosis' is often used as a shorthand for 'early-stage diagnosis', which has historically been the outcome in early diagnosis studies. However, alternative indicators based on electronic health records are increasingly being used in early diagnosis studies. Some of these indicators relate to the promptness of diagnosis following clinical presentation, or the health services patients accessed first [6].

In England, cancer surveillance statistics are published on Public Health England's 'CancerData' dashboard for each of the 209 local healthcare commissioning bodies (Clinical Commissioning Groups – CCGs) [7]. These include the percentage of patients diagnosed with localised tumours (TNM Stage I/II), the percentage diagnosed following emergency admission or referral, and statistics on adherence to targets for patient waiting times.

Surveillance in England was only initiated in 2016, and optimal implementation of the different possible indicators in surveillance has not been extensively researched. Information on the association between each indicator and short-term mortality and survival will help analysts interpret the indicators, and identify those which are timely measures of progress in raising survival from cancer.

In this study, we report a systematic literature review and data analysis. Our aim is to identify early diagnosis indicators and evaluate the association between each indicator and short-term mortality and survival. We then discuss the implications of our findings for surveillance.

2. Methods

We conducted a systematic literature search to identify indicators of early cancer diagnosis in population-based data sources and evaluate their association with short-term mortality and survival. The association between each indicator and (i) risk of death within 30 days following diagnosis ("30-day mortality") (ii) net survival at one and five

* Corresponding author.
E-mail address: patrick.muller@lshtm.ac.uk (P. Muller).

https://doi.org/10.1016/j.canep.2018.07.010
Received 20 March 2018; Received in revised form 20 June 2018; Accepted 16 July 2018
Available online 25 July 2018

years was then analysed for patients diagnosed in 2009–2013 in England with either colorectal cancer, non-small cell lung cancer (NSCLC) or ovarian cancer, malignancies frequently diagnosed at advanced stage [8].

2.1. Literature review

2.1.1. PubMed search

Journal articles published between August 2007 and August 2017 were examined [9]. Articles that contained each of three keyword elements (cancer, early diagnosis, population-based) in the title or the abstract were retained (Appendix A in Supplementary data).

2.1.2. Google.com search

Google was searched using the same keywords on 25 August 2017. The first 20 hyperlinks returned were exhaustively searched for relevant journal articles or reports, and any found were retained.

2.1.3. Document selection strategy

The abstract of each retained article was read. If it reported new statistics or methods for generating statistics based on an indicator of early diagnosis from population-based data sources (inclusion criterion 1, Appendix A in Supplementary data), the article's full text was read.

The summary, introduction and conclusion of retained reports were read. Those meeting the above inclusion criterion were identified. Relevant portions of the full text of these reports were then read.

Details regarding the early diagnosis indicators and their association with any measure related to survival (complications, mortality, survival, life expectancy, or cure (inclusion criterion 2, Appendix A in Supplementary data)) were extracted.

2.2. Data analysis

2.2.1. Patient cohort

Cancer registrations were obtained from the Office for National Statistics (ONS) for adults aged 15–99 years, diagnosed with colorectal cancer, NSCLC or ovarian cancer in England in 2006–2013 (ICD-10 codes C18-20, C21.8, C33-34 and C56-C57.7 [10]). Follow-up was complete up to 31 December 2014. Registrations were linked to datasets from the National Bowel and Lung Cancer Audits, and to the Routes to Diagnosis dataset [11] using the patient's NHS number and postcode. These datasets were used to complete information on stage at diagnosis [12].

2.2.2. Data analysis

Thirty-day mortality and one- and five-year net survival were estimated by agegroup (15–59, 60–79, 80–99 years). Net survival estimates were obtained using Pohar Perme's unbiased estimator [13] and the period approach applied to follow-up data during 2009–2013 for patients diagnosed during 2006–2013 [14] (details in Appendix B in Supplementary data).

To avoid unstable sub-group estimates, one-year survival was only estimated if, in the period 2009–2013, at least 25 patients were diagnosed and five deaths occurred within the first year after diagnosis. Five-year survival was estimated if at least 15 patients were alive at one year after diagnosis and five deaths occurred in the second to fifth years. Missing data was either included in a separate category, or excluded (complete-case analysis).

3. Results

3.1. Literature search

The PubMed search returned 154 articles (Fig. 1), 19 of which presented new statistics or methods for generating statistics for an early diagnosis indicator. The Google search returned five reports and six articles also meeting that criterion.

Three early diagnosis indicators were identified in these 30 documents: stage at diagnosis (21 documents), emergency admission or emergency presentation (five), and interval from first symptom to diagnosis (eight).

Four documents contained information on survival and two on mortality.

3.1.1. Indicator 1: Stage at diagnosis

3.1.1.1. Definition and description. Stage was the sole indicator used in 18 documents (Table 1), and was one of several indicators in a further three. Typically the TNM classification system was used, directly or using ordinal stage (I–IV). Dukes' stage for colorectal cancer [16] and tumour thickness for melanoma [23] were also used. Stage was frequently dichotomised into 'early' (localised, stage I or II, non-metastatic) vs. 'late' (advanced, stage III or IV, metastatic) [16,22]. The *CancerData* dashboard uses a binary indicator for whether the patient has a record of stage I/II disease, and presents this as a percentage of all patients (including those without a recorded stage) [37]. Other studies imputed stage information [15,19], or analysed missing stage as a separate group [35]. Average stage at diagnosis varied by tumour site [38], histological type for ovarian cancer [20], and by subsite for colorectal cancer [33].

3.1.1.2. Association with patient survival. Early-stage disease was associated with higher survival (net, relative or unadjusted). Women diagnosed with breast cancer during 2000–2007 in northeast England had one-year net survival over 90% at stages I-III, but only 50% at stage IV [15]. Five-year survival was high for women diagnosed at stage I/II, substantially lower for stage III and lower still for stage IV. There was a similar pattern for patients diagnosed in England during 1996–2002 with colorectal cancer: five-year survival was 93.2% for the patients diagnosed with Dukes' A, 47.7% for Dukes' C, and 6.6% for stage D [35].

3.1.2. Indicator 2: Emergency admission or presentation

3.1.2.1. Definition and description. Emergency admission or presentation was mentioned in five documents (Table 1). This is defined as an admission coded as 'emergency' and/or 'accident & emergency' [16], or a route to diagnosis via the Accident & Emergency department or via an emergency referral or transfer [11].

3.1.2.2. Association with short-term mortality and survival. Elliss-Brookes et al. [11] found an association between emergency presentation and 1-year relative survival, noting "the substantially lower relative survival in emergency compared to non-emergency routes indicates that this distinction is of high clinical significance".

3.1.3. Indicator 3: Interval from first symptoms to diagnosis

3.1.3.1. Definition and description. Time from first cancer-relevant symptom to referral or diagnosis was used in eight documents (Table 1).

The interval start was the time the patient first noticed symptoms [34,39,40] or time of presentation with symptoms to the GP [17,18,42]. The relevance of a symptom to cancer was decided by the GP [42], specialist clinician review [18], or by reference to external standards [40].

The end-point was cancer diagnosis or referral to secondary care. One study defined two intervals: the patient interval (time from symptom onset to first clinical presentation), and the primary care interval (time from first clinical presentation to specialist referral) [39].

3.1.3.2. Association with short-term mortality and survival. The association between the intervals and survival varied by cancer. One study of childhood acute lymphoblastic leukaemia (ALL) found a

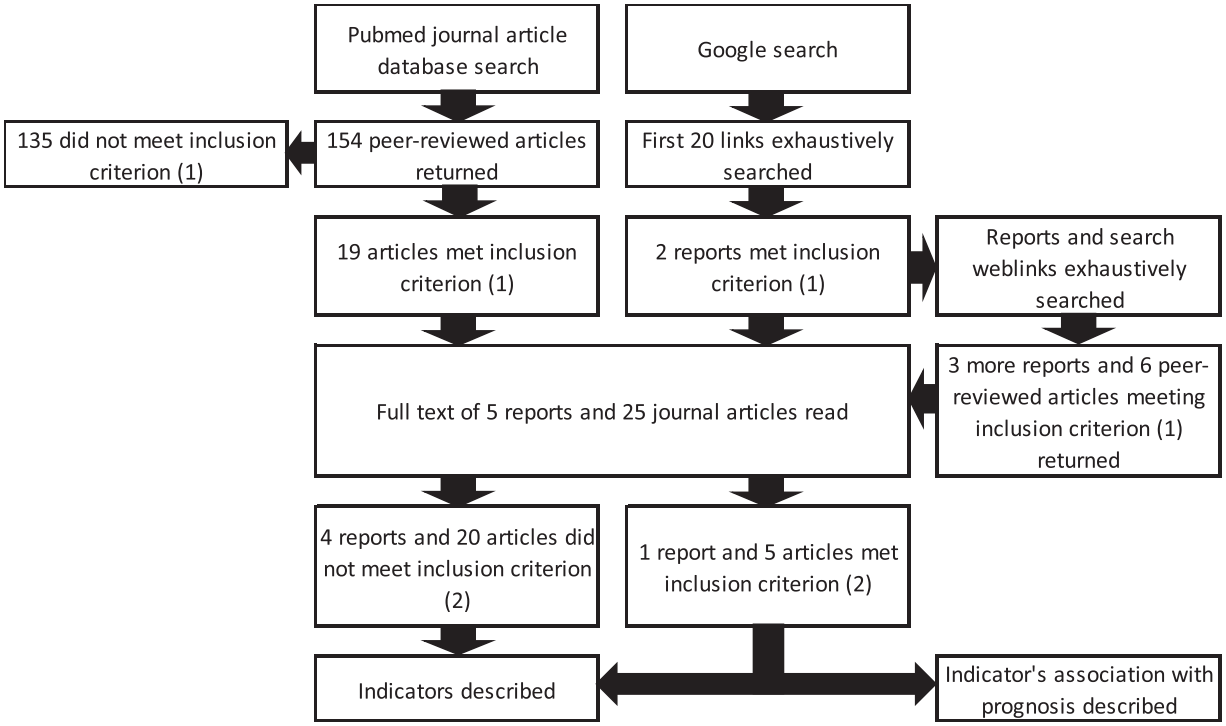


Fig. 1. Search strategy with number of reports and journal articles which reported new early diagnosis statistics or methods for generating these (inclusion criterion (1)), and number reporting the association between an indicator and a measure related to survival (inclusion criterion (2)).

prolonged interval from presentation to diagnosis was associated with longer event-free survival, although this was attributed to confounding from disease biology [18]. Another study found a 'U-shaped' curve between interval and odds of death within five years for five common cancers, with higher odds for patients with the shortest and longest intervals [42]. The high odds of death amongst patients with short intervals was attributed to confounding, arising because of GPs expediting diagnosis for patients with high-risk symptoms.

In one survey of expert judgement for 21 common cancers [43,44] there was consensus that expedited diagnosis brings mortality reductions for 11 cancers. They were undecided for seven cancers, and disagreed that expedited diagnosis conferred any mortality benefit for three.

3.2. Data analysis

We analysed the association between stage at diagnosis and emergency presentation with 30-day mortality and survival for 160,617 colorectal, 170,425 non-small cell lung (NSCLC), and 24,450 ovarian cancer patients (Table 2). Data was not available to examine the interval from first symptoms to diagnosis.

3.2.1. Stage at diagnosis: association with 30-day mortality and net survival

Stage was missing for a large proportion of patients in the linked datasets we analysed (16.1–36.4% of patients for the three cancers). Colorectal and NSCLC patients aged 60–79 were most likely to be diagnosed at stages I or II (43% and 22% respectively), but there were not substantial differences between age groups (Table 2). By contrast, ovarian cancer patients aged 15–59 were considerably more likely to be diagnosed at stage I/II than those aged 80–99 (48% vs 22%).

Risk of 30-day mortality was higher at higher stages of disease, for all cancers and age groups, with the exception of colorectal cancer where mortality risk plateaued at stages II-III (Table 2). Thirty-day mortality was considerably higher for stage IV patients than stage III patients: six-times higher for colorectal cancer and three-times higher for NSCLC and ovarian cancer patients. NSCLC and ovarian cancer patients with missing stage had even higher mortality than stage IV patients (37.2% vs. 22.8% for NSCLC; 16.9% vs.12% for ovarian cancer) whereas mortality for colorectal cancer patients with missing stage was between that of patients diagnosed at stages III and IV.

One-year colorectal cancer survival was similar for patients diagnosed at stages I-III (9.6% difference between stages I and III) but markedly lower for stage IV patients (Table 3, Fig. 2). There was no such plateau in five-year survival (32.1% difference between stages I and III). For NSCLC and ovarian cancer, incremental differences in stage category (I vs II, II vs III, III vs IV) were associated with significantly lower one-year and five-year survival; no plateau was evident. Patients missing stage had low survival, typically between the survival of patients with stage III and stage IV disease (Tables 3 and 4).

3.2.2. Emergency presentation: association with 30-day mortality and net survival

Twenty-three percent of colorectal, 35.6% of NSCLC, and 30.7% of ovarian cancer patients were diagnosed following emergency presentation (Appendix C in Supplementary data). Emergency presentation risk was greater for older patients diagnosed with NSCLC and ovarian cancer, whilst for colorectal cancer it was most common for patients aged 15–59 and 80–99 (Table 2).

Emergency presentation was associated with 1.9–2.9 times higher 30-day mortality (Table 2) and lower one-year net survival (Table 3): 50.7% for colorectal cancer compared with 75.9% survival for all routes combined; 14.1% compared to 32.6% for NSCLC; and 43.7% compared to 68.2% for ovarian cancer. Differences were greater for older patients. Similar patterns were observed for five-year survival (Table 4).

A small proportion (1.8–2.4%) of patients could not be assigned a route to diagnosis. Applying the assumption that these were all non-emergency presentations resulted in small changes in net survival (typically < 1% and never > 2%) indicating that these results are not sensitive to missing data (Appendix D in Supplementary data).

Table 1
Cancers investigated, early diagnosis indicators and survival measures used in the documents read in full.

Reference		Early diagnosis indicator	Survival measure	Document type	Cancer(s)	Study population
Ahrensberg et al. [17]	~	Interval from first presentation to diagnosis	–	Journal article	Childhood cancers/benign tumours of the Central nervous system (CNS)	Population-based analysis of Danish children diagnosed in 2007–2010
Aitken et al. [23]	^	Tumour thickness (one component of TNM stage)	–	Journal article	Melanoma	Population-based analysis of Australian people aged 20–75 diagnosed in 2000–2003
Castanon et al. [24]	^	Stage†	–	Journal article	Cervix	Population-based analysis of English and Welsh women aged 30–69 diagnosed in 1990–2014
Chorley et al. [26]	^	Stage	–	Journal article	Cervix, breast, colorectal	Survey of English adults aged 50–70 conducted in 2015
Ciocan et al. [25]	^	Stage	–	Journal article	Melanoma	Population-based analysis of French people diagnosed with melanoma in 2004–2008
Durbec et al. [27]	^	Tumour thickness (one component of TNM stage)	–	Journal article	Melanoma	Population-based analysis of French people diagnosed in 2004
Elliss-Brookes et al. [45]	¥	Emergency presentation	Relative survival from diagnosis	Journal article	15 common cancers	Population-based analysis of English people diagnosed in 2006–2008
Eskesen et al. [28]	^	Stage	–	Journal article	Liver	Population-based analysis of Norwegian people diagnosed in 2000–2009
Ess et al. [29]	^	Stage	–	Journal article	Breast	Representative sample study of Swiss women diagnosed in 2003–2005
Greenlee et al. [30]	^	Stage (Local vs distant)	–	Journal article	Larynx, oral cavity, melanoma, breast, prostate, corpus uteri, cervix, bladder, colorectum, esophagus, stomach, kidney	Population-based analysis of people in the USA diagnosed in 1997–2000
Gupta et al. [18]	~	Interval from first presentation to diagnosis	Event-free survival from diagnosis	Journal article	Acute lymphoblastic leukaemia	Population-based analysis of Canadian children diagnosed in 1995–2011
Hamilton et al. [43]	~	Interval from first symptoms to diagnosis	Mortality amongst symptomatic patients	Journal article	21 common cancers	Qualitative survey of 22 UK and Danish experts conducted in 2014
Hreinsson et al. [31]	^	Stage (I-IV, non-metastatic vs metastatic)	–	Journal article	Colorectal	Population-based analysis of Irish people diagnosed 2008–2011
Independent Cancer Taskforce [38]	^ ¥ ~	Stage, emergency presentation, interval from first recognition of symptoms to diagnosis	–	Report	All cancers	None, specification of an early diagnosis statistic
Laudicella et al. [19]	^	Stage I/II vs III/IV	–	Journal article	Colorectal, prostate, lung, breast	Population-based analysis of English women diagnosed in 2001–2010
Li et al. [15]	^	Stage	Net cancer survival from diagnosis	Journal article	Breast	Population-based analysis of English women diagnosed in 2000–2007
Lurie et al. [32]	^	Stage	–	Journal article	Ovarian	Population-based analysis of women in the USA diagnosed in 1993–1997
Lurie et al. [20]	^	Stage	–	Journal article	Ovarian	Population-based analysis of women diagnosed in the USA in 1993–2008
Lyratzopoulos et al. [49]	~	Interval from first recognition of symptoms to referral, broken into patient and primary care intervals	–	Journal article	28 adult cancers	Population-based study of English patients presenting in primary care in 2009–2010
Murage et al. [16]	^ ¥	Stage (Duke's AB vs CD), Emergency admission		Journal article	Colorectal	Population-based analysis of Scottish patients diagnosed in 1998–2011
National Cancer Intelligence Network [35]	^	Stage	Net cancer survival from diagnosis	Report	Colorectal	Population-based study of adults diagnosed in England in 1996–2002
Public Health England [36]	¥	Emergency presentation	–	Report	All invasive malignancies	None, specification of an early diagnosis statistic
Public Health England [37]	^	Stage (I, II vs III, IV, missing)	–	Report	Breast, prostate, colorectal, lung, baldder, kidney, ovary, uterus, non-Hodgkin lymphomas, melanoma	None, specification of an early diagnosis statistic
Rubin et al. [41]	¥	Emergency presentation	–	Journal article	All cancers	Mixed-methods (interviews and data analysis) study of the English GP practices in the period 2009–2013
Sala et al. [21]	^	Stage	–	Journal article	Breast	Population-based analysis of Spanish women aged 50-69 who attended screening in 1995–2010

(*continued on next page*)

Table 1 (*continued*)

Reference		Early diagnosis indicator	Survival measure	Document type	Cancer(s)	Study population
Sankaranarayanan et al. [33]	^	Stage (localised, in-situ vs regional, distant)	–	Journal article	Colorectal	Population-based analysis of people in the USA diagnosed 1998–2003
Torring et al. [42]	~	Interval from first presentation to diagnosis	Mortality risk within 5 years of diagnosis	Journal article	Colorectal, lung, melanoma, breast, prostate	Population-based analysis of Danish adults diagnosed in 2004–2005
Tracey et al. [22]	^	Stage (localised vs advanced)	–	Journal article	[All cancers]	Population-based analysis for Australian people diagnosed in 1980–2008
Walter et al. [40]	~	Interval from first recognition of symptoms to diagnosis	–	Journal article	Lung	Population-based study of English people referred to secondary care in 2010–2012
World Health Organisation [34]	^ ~	Stage, interval from first recognition of symptoms to diagnosis	–	Report	All cancers	None, specification of an early diagnosis statistic

†"Stage" = "Stage of disease at diagnosis".
^Stage.
¥Emergency presentation.
~Interval from firsy symptoms to diagnosis.

3.2.3. Stage at diagnosis and emergency presentation: joint association with 30-day mortality and net survival

Patients diagnosed following emergency presentation were more likely to be diagnosed at stages III or IV or have stage missing (Table 2). Emergency presentation was associated with higher 30-day mortality and lower one- and five-year net survival for patients at each stage (Tables 2–4). Survival differences between emergency and non-emergency colorectal cancer patients increased after the first year of follow up (Appendix E.1 in Supplementary data). By contrast, survival for emergency and non-emergency NSCLC patients converged to a 'floor' by the fifth year of follow up (Appendix E.2 in Supplementary data). Patients diagnosed following emergency presentation with missing stage had extremely high mortality: 26.4–51.0% died within 30 days following diagnosis (Table 2).

4. Discussion

Stage at diagnosis, emergency admission or presentation, and interval from first symptoms to diagnosis are commonly used indicators of early diagnosis. However, in the literature only stage and emergency diagnosis have a straightforward relationship with patient survival.

Our data analysis showed that emergency presentation and stage are independently associated with higher 30-day mortality and lower survival from colorectal, NSCLC and ovarian cancer in England. Patients without a recorded stage in population-based datasets had extremely high 30-day mortality and lower five-year survival.

4.1. Association between stage and survival

One-year survival from colorectal and breast cancers plateaued at stages I-III and was markedly lower at stage IV, whereas NSCLC and ovarian cancer displayed no such plateau. Five-year survival did not plateau at any stages for any cancer: each incremental increase in stage was associated with substantially lower survival. Granular information on stage at diagnosis, as opposed binary groupings, is therefore useful for monitoring progress in efforts to raise medium-term survival, although certain binary stage groupings may produce statistics which are strongly associated with short-term survival.

4.2. Association between emergency presentation and short-term mortality and survival

We found that emergency presentation was associated with higher 30-day mortality and lower medium-term survival for patients at every age and stage disease, consistent with other studies [44,45]. This indicator is therefore a proxy for other factors which independently determine survival, and is a valuable complimentary prognostic indicator to stage. However, more work is needed to understand why it is independently associated with survival.

4.3. Association between interval from first symptoms to diagnosis and short-term mortality and survival

Shorter intervals from first symptoms to diagnosis were not consistently associated with improved survival in the literature we examined. Other reviews concur. Neal et al. found instances of contradiction between studies on a given cancer on whether reducing the diagnostic, referral, or treatment interval was associated with higher survival or reduced mortality [46]. Hamilton et al found there was no consensus between experts that expediting symptomatic diagnosis conferred a mortality benefit for many common cancers [43].

These inconsistent findings may be partly explained by confounding by tumour aggressiveness and stage. The 'waiting times paradox' of the shortest intervals being associated with poor survival [42,47] is also likely to be partially attributable to confounding by these tumour factors [48]: Stage and aggressiveness may determine both type of first

Table 2
Number of patients (%), and risk of death within 30 days following diagnosis (Mortality) by age group, route, cancer, and stage at diagnosis, England 2009–2013.

	All ages						Ages 15–59					
	All routes			Emergency presentation only			All routes			Emergency presentation only		
	N	(%)	*Mortalty %*	N	(%)	*Mortalty %*	N	(%)	*Mortalty %*	N	(%)	*Mortalty %*
Colorectal cancer												
I	17,180	(10.9)	*0.4*	1,142	(3.1)	*3.4*	2,752	(10.7)	*0.1*	314	(5.2)	*0.6*
II	27,579	(17.5)	*2.0*	5,158	(13.8)	*7.4*	4,056	(15.7)	*0.2*	843	(13.8)	*0.8*
III	29,680	(18.8)	*1.8*	5,788	(15.5)	*6.3*	5,528	(21.4)	*0.2*	1,064	(17.5)	*0.8*
IV	33,215	(21.1)	*11.2*	11,059	(29.7)	*23.9*	6,300	(24.4)	*4.9*	1,965	(32.2)	*10.1*
Missing	50,125	(31.8)	*10.6*	14,138	(37.9)	*26.4*	7,151	(27.7)	*3.1*	1,909	(31.3)	*7.1*
Total	157,779	(100.0)	*6.5*	37,285	(100.0)	*19.2*	25,787	(100.0)	*2.1*	6,095	(100.0)	*5.8*
Non-small cell lung cancer												
I	19,018	(11.4)	*2.3*	3,440	(5.7)	*7.9*	2,053	(9.6)	*0.4*	233	(3.6)	*1.3*
II	10,921	(6.6)	*3.6*	2,127	(3.5)	*11.2*	1,268	(5.9)	*1.0*	165	(2.5)	*4.2*
III	34,151	(20.5)	*7.8*	8,316	(13.7)	*20.4*	4,466	(20.9)	*4.1*	818	(12.6)	*13.2*
IV	75,506	(45.4)	*22.8*	33,168	(54.7)	*36.8*	10,708	(50.1)	*14.5*	4,188	(64.5)	*24.7*
Missing	26,817	(16.1)	*37.2*	13,581	(22.4)	*51.0*	2,865	(13.4)	*22.3*	1,086	(16.7)	*39.7*
Total	166,413	(100.0)	*18.5*	60,632	(100.0)	*35.2*	21,360	(100.0)	*11.2*	6,490	(100.0)	*24.4*
Ovarian cancer												
I	3,989	(16.7)	*0.6*	472	(6.3)	*2.3*	1,983	(27.3)	*0.1*	282	(16.5)	*0.7*
II	1173	(4.9)	*1.9*	147	(2.0)	*8.2*	449	(6.2)	*0.7*	56	(3.3)	*3.6*
III	5,745	(24.0)	*4.0*	1,555	(20.7)	*10.0*	1,672	(23.0)	*1.3*	423	(24.7)	*3.3*
IV	4,288	(17.9)	*12.0*	1,858	(24.8)	*20.5*	956	(13.2)	*4.5*	360	(21.0)	*8.9*
Missing	8,712	(36.4)	*16.9*	3,471	(46.3)	*30.7*	2,206	(30.4)	*5.3*	592	(34.6)	*12.5*
Total	23,907	(100.0)	*9.5*	7,503	(100.0)	*21.6*	7,266	(100.0)	*2.6*	1,713	(100.0)	*7.2*

	Ages 60–79						Ages 80–99					
	All routes			Emergency presentation only			All routes			Emergency presentation only		
	N	(%)	*Mortalty %*	N	(%)	*Mortalty %*	N	(%)	*Mortalty %*	N	(%)	*Mortalty %*
Colorectal cancer												
I	11,173	(12.5)	*0.2*	420	(2.6)	*1.2*	3,255	(7.6)	*1.6*	408	(2.7)	*7.8*
II	16,056	(18.0)	*1.3*	2,444	(15.0)	*5.7*	7,467	(17.5)	*4.5*	1,871	(12.6)	*12.5*
III	17,479	(19.6)	*1.3*	2,833	(17.3)	*5.2*	6,673	(15.7)	*4.4*	1,891	(12.7)	*10.9*
IV	18,283	(20.5)	*9.4*	5,329	(32.6)	*21.5*	8,632	(20.3)	*19.8*	3,765	(25.4)	*34.4*
Missing	26,380	(29.5)	*7.1*	5,312	(32.5)	*22.4*	16,594	(38.9)	*19.4*	6,917	(46.6)	*34.8*
Total	89,371	(100.0)	*4.5*	16,338	(100.0)	*16.1*	42,621	(100.0)	*13.2*	14,852	(100.0)	*28.1*
Non-small cell lung cancer												
I	12,303	(12.1)	*1.9*	1,769	(5.4)	*6.6*	4,662	(10.7)	*4.3*	1,438	(6.7)	*10.5*
II	6,968	(6.9)	*3.1*	1,095	(3.4)	*11.1*	2,685	(6.1)	*6.3*	867	(4.0)	*12.7*
III	21,795	(21.5)	*6.6*	4,587	(14.1)	*18.4*	7,890	(18.0)	*13.1*	2,911	(13.5)	*25.8*
IV	46,260	(45.7)	*21.5*	18,755	(57.6)	*36.0*	18,538	(42.4)	*31.0*	10,225	(47.3)	*43.4*
Missing	14,004	(13.8)	*35.2*	6,327	(19.4)	*51.3*	9,948	(22.8)	*44.4*	6,168	(28.5)	*52.7*
Total	101,330	(100.0)	*16.5*	32,533	(100.0)	*34.0*	43,723	(100.0)	*26.5*	21,609	(100.0)	*40.3*
Ovarian cancer												
I	1,688	(13.7)	*0.4*	147	(4.0)	*2.0*	318	(7.3)	*4.4*	43	(2.0)	*14.0*
II	591	(4.8)	*1.4*	65	(1.8)	*6.2*	133	(3.0)	*8.3*	26	(1.2)	*23.1*
III	3,329	(27.1)	*3.4*	859	(23.6)	*8.5*	744	(17.1)	*12.5*	273	(12.7)	*24.9*
IV	2,432	(19.8)	*9.4*	992	(27.3)	*16.4*	900	(20.6)	*26.8*	506	(23.5)	*36.8*
Missing	4,240	(34.5)	*13.8*	1,572	(43.2)	*26.1*	2,266	(52.0)	*34.0*	1,307	(60.6)	*44.3*
Total	12,280	(100.0)	*7.7*	3,635	(100.0)	*18.0*	4,361	(100.0)	*25.9*	2,155	(100.0)	*39.2*

Table 3
One-year net survival (and 95% confidence interval) by age group, route, cancer, and stage at diagnosis, England 2009–2013.

	Ages 15–59		Ages 60–79		Ages 80–99	
	All routes	Emergency presentation	All routes	Emergency presentation	All routes	Emergency presentation
Colorectal cancer						
I	99.3 (98.9, 99.7)	98.3 (96.7, 100.0)	98.8 (98.4, 99.1)	90.5 (87.2, 93.8)	93.7 (92.3, 95.1)	67.6 (61.9, 73.3)
II	97.8 (97.2, 98.3)	95.5 (93.9, 97.0)	95.2 (94.8, 95.6)	86.6 (85.1, 88.2)	87.7 (86.7, 88.8)	72.0 (69.5, 74.5)
III	95.4 (94.8, 96.0)	90.0 (88.0, 91.9)	90.6 (90.1, 91.1)	78.3 (76.7, 80.0)	76.2 (74.9, 77.5)	58.4 (55.7, 61.0)
IV	62.9 (61.7, 64.1)	48.9 (46.6, 51.2)	50.0 (49.3, 50.8)	31.2 (30.0, 32.5)	27.6 (26.6, 28.6)	15.2 (14.0, 16.5)
Missing	86.5 (85.7, 87.2)	75.6 (73.8, 77.5)	79.6 (79.1, 80.1)	52.1 (50.8, 53.4)	52.5 (51.7, 53.3)	30.2 (29.1, 31.3)
Total	85.6 (85.1, 86.0)	73.2 (72.0, 74.3)	80.7 (80.4, 81.0)	55.8 (55.0, 56.5)	59.9 (59.4, 60.4)	35.9 (35.0, 36.7)
Non-small cell lung cancer						
I	92.9 (91.8, 94.1)	84.9 (80.1, 89.8)	85.0 (84.3, 85.7)	67.0 (64.6, 69.4)	70.2 (68.6, 71.8)	49.2 (46.2, 52.2)
II	81.4 (79.1, 83.6)	70.7 (63.3, 78.0)	69.0 (67.8, 70.2)	45.9 (42.7, 49.0)	48.0 (45.9, 50.2)	30.5 (27.0, 34.0)
III	54.3 (52.8, 55.8)	34.6 (31.3, 37.9)	43.2 (42.5, 43.8)	23.5 (22.3, 24.8)	26.8 (25.8, 27.9)	14.6 (13.2, 16.0)
IV	24.0 (23.2, 24.9)	15.0 (13.9, 16.1)	16.9 (16.5, 17.2)	7.6 (7.3, 8.0)	10.0 (9.5, 10.5)	5.2 (4.8, 5.7)
Missing	39.4 (37.8, 41.1)	19.1 (17.0, 21.3)	24.4 (23.7, 25.0)	9.4 (8.8, 10.1)	12.3 (11.7, 12.9)	6.8 (6.2, 7.4)
Total	42.4 (41.7, 43.0)	22.1 (21.1, 23.1)	35.2 (34.9, 35.5)	14.7 (14.3, 15.1)	22.0 (21.6, 22.4)	10.7 (10.3, 11.2)
Ovarian cancer						
I	98.1 (97.4, 98.7)	96.0 (93.6, 98.4)	97.1 (96.1, 98.1)	90.3 (85.1, 95.5)	93.2 (89.2, 97.3)	64.4 (48.4, 80.4)
II	92.3 (89.7, 94.8)	80.4 (70.1, 90.8)	90.3 (87.6, 93.0)	83.7 (74.2, 93.2)	73.0 (64.3, 81.8)	36.4 (16.9, 56.0)
III	87.7 (86.0, 89.3)	79.7 (75.7, 83.8)	76.3 (74.8, 77.9)	59.7 (56.2, 63.1)	49.3 (45.3, 53.3)	27.2 (21.4, 33.0)
IV	72.2 (69.3, 75.1)	64.9 (59.9, 69.9)	57.3 (55.3, 59.4)	43.6 (40.4, 46.8)	21.4 (18.5, 24.2)	11.2 (8.3, 14.1)
Missing	82.0 (80.4, 83.5)	65.0 (61.3, 68.7)	58.4 (56.9, 59.8)	36.6 (34.2, 38.9)	22.7 (21.0, 24.5)	11.9 (10.1, 13.6)
Total	86.9 (86.1, 87.6)	73.9 (71.8, 76.0)	69.5 (68.7, 70.4)	46.8 (45.1, 48.4)	33.1 (31.7, 34.6)	14.8 (13.2, 16.3)

presenting symptoms (in turn determining interval length) and patient survival. We found evidence suggesting this in the literature: type of first symptoms was associated with the length of interval for childhood CNS [17], lung cancer [40], and ovarian cancer [20].

4.4. Monitoring performance using early diagnosis indicators

Monitoring of early-stage diagnosis is England is currently conducted using the percentage of patients diagnosed at stage I or II. However, we have shown that binary groupings of stage lose information which is predictive of medium-term survival. Numerical average stage (1–4) might provide a simple alternative measure that is more strongly associated with medium-term survival. Within a modelling framework ordered logistic regression with 4-category stage could be used instead of logistic regression with a binary stage indicator.

Emergency presentation is associated with advanced stage, and higher mortality and lower survival for patients at each stage. Data on emergency presentation could therefore be combined with stage information to generate a more informative prognostic index. Patients newly diagnosed through a given route at a given stage could be assigned a score which is the average survival of patients previously diagnosed with the same combination of indicators. For example, if a patient were diagnosed via emergency presentation at stage II disease, and previous 1-year survival for patients with these attributes was 80%, then that patient would be assigned a score of 80. The average score of the patient population could be then be used for monitoring and comparisons.

Our results don't support the use of 'average diagnostic interval length' statistics for benchmarking and performance management. This is because the very shortest intervals are associated with poorer survival (due to confounding), so short intervals are not necessarily indicative of success in early diagnosis. However, it would be worthwhile

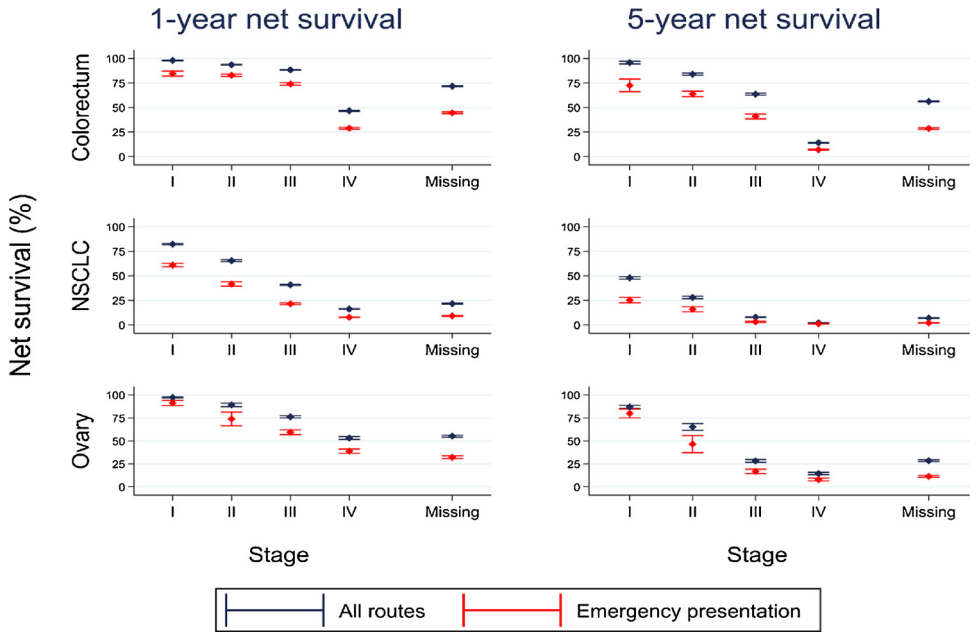


Fig. 2. Net survival (with 95% confidence intervals (CI)) at one and five years by cancer, stage, and route to diagnosis, England.2009–2013.

Table 4
Five-year net survival (and 95% confidence interval) by age group, route, cancer, and stage at diagnosis, England 2009–2013.

	Ages 15–59		Ages 60–79		Ages 80–99	
	All routes	Emergency presentation	All routes	Emergency presentation	All routes	Emergency presentation
Colorectum						
I	94.5 (92.6, 96.4)	93.4 (87.2, 99.5)	97.1 (95.7, 98.4)	74.4 (65.8, 82.9)	91.8 (86.2, 97.4)	58.4 (44.9, 71.9)
II	87.3 (85.6, 89.1)	78.4 (73.7, 83.2)	86.2 (85.0, 87.4)	68.5 (65.3, 71.8)	77.2 (73.7, 80.6)	50.6 (44.1, 57.0)
III	71.3 (69.4, 73.1)	57.5 (52.7, 62.2)	66.9 (65.7, 68.1)	42.3 (39.3, 45.4)	47.9 (44.9, 50.9)	29.7 (24.1, 35.2)
IV	19.7 (18.4, 20.9)	12.9 (11.0, 14.8)	15.3 (14.6, 16.0)	7.2 (6.3, 8.1)	7.3 (6.4, 8.3)	3.8 (2.6, 5.0)
Missing	68.9 (67.9, 69.8)	55.4 (53.3, 57.5)	64.2 (63.6, 64.8)	33.5 (32.3, 34.7)	37.7 (36.7, 38.7)	17.6 (16.4, 18.8)
Overall	63.2 (62.6, 63.9)	47.4 (45.9, 48.8)	62.0 (61.6, 62.5)	33.0 (32.2, 33.8)	42.9 (42.1, 43.8)	20.6 (19.5, 21.6)
NSCLC						
I	67.7 (65.0, 70.5)	52.6 (44.4, 60.7)	50.8 (49.4, 52.3)	29.7 (26.2, 33.1)	29.8 (26.8, 32.8)	13.3 (8.4, 18.1)
II	46.3 (42.7, 50.0)	43.1 (33.4, 52.7)	29.9 (28.1, 31.6)	18.5 (15.1, 21.9)	12.9 (10.0, 15.8)	6.0 (1.7, 10.3)
III	13.1 (11.9, 14.3)	7.5 (5.5, 9.6)	8.3 (7.8, 8.8)	3.3 (2.6, 4.0)	3.5 (2.8, 4.3)	1.2 (0.4, 2.0)
IV	3.9 (3.4, 4.3)	2.6 (2.0, 3.3)	1.8 (1.6, 2.0)	0.8 (0.6, 1.0)	1.3 (1.0, 1.7)	1.1 (0.6, 1.5)
Missing	17.9 (16.8, 19.0)	6.8 (5.6, 8.0)	7.9 (7.5, 8.3)	2.2 (1.9, 2.5)	2.5 (2.1, 2.8)	0.8 (0.5, 1.0)
Overall	16.6 (16.0, 17.1)	6.9 (6.2, 7.5)	11.5 (11.3, 11.8)	3.5 (3.2, 3.7)	5.3 (4.9, 5.6)	1.7 (1.4, 2.1)
Ovarian cancer						
I	88.3 (86.4, 90.2)	86.8 (82.1, 91.5)	86.5 (83.8, 89.3)	72.9 (63.2, 82.5)	81.6 (70.0, 93.1)	53.4 (26.5, 80.3)
II	74.7 (70.0, 79.4)	55.9 (42.4, 69.3)	63.1 (57.5, 68.7)	49.7 (34.6, 64.9)	37.2 (23.3, 51.1)	15.8 (-0.2, 31.8)
III	41.5 (38.4, 44.5)	31.5 (25.8, 37.2)	24.9 (22.9, 26.9)	14.4 (11.4, 17.4)	11.4 (7.4, 15.5)	3.6 (0.5, 6.6)
IV	26.6 (23.4, 29.8)	18.4 (13.5, 23.3)	13.8 (12.1, 15.5)	8.2 (6.1, 10.2)	3.6 (1.7, 5.4)	0.0 (0.0, 0.1)
Missing	55.2 (53.4, 57.0)	33.5 (30.2, 36.7)	25.1 (23.9, 26.2)	9.5 (8.2, 10.8)	8.1 (6.9, 9.4)	3.3 (2.2, 4.4)
Overall	57.9 (56.7, 59.1)	38.9 (36.5, 41.3)	31.7 (30.8, 32.6)	13.1 (12.0, 14.3)	13.4 (12.0, 14.8)	3.8 (2.7, 4.8)

to monitor whether reductions in average diagnostic intervals in response to an intervention in an area are associated with changes in the stage distribution or survival, to evaluate the effectiveness of the intervention.

Further work is also needed to identify alternative statistics based on the interval or similar quantities which are useful for surveillance. Alternative measures could include statistics on 'missed opportunities' for prompt symptomatic diagnosis: Lyratzopoulos et al. have described how these can occur [49], and Renzi et al identified instances of these for colorectal cancer [50].

We found that anatomical site of origin was strongly associated with probability of early diagnosis, regardless of the indicator used. The distribution of cancers should therefore be accounted for in performance comparisons, either through standardisation a modelling approach, to reduce bias from case-mix differences.

4.5. Interpreting missing stage information

In the English datasets we examined 16–36% of patients had no recorded stage. Compared to patients with a recorded stage, these patients had very high risk of death shortly following diagnosis and lower medium-term survival.

There are likely to be two different reasons why patients do not have a recorded stage. For most patients missing stage, it may be for administrative (non-clinical) reasons. These patients would have a similar stage distribution and survival to those with recorded stage. For a minority, it may have not been recorded *because* the patient was acutely unwell or had very poor prognosis at the time of first presentation. These patients would have more advanced disease, and many would die shortly after diagnosis. This hypothesis, that the majority of patients missing stage have typical stage and survival, and a minority have more advanced disease and very poor survival, explains the heterogeneity in 30-day mortality and survival we observed. It is also consistent with results from the study by Barclay et al suggesting that the stage distribution of these patients is slightly skewed towards later stages [51].

Our findings suggest that patients missing stage should be included in surveillance to avoid bias. This could be done using multiple imputation for missing data [52], using a 'missing stage and died shortly following diagnosis' percentage, or by applying expected survival statistics or model-based scores.

4.6. Strengths and limitations

We conducted a comprehensive joint analysis of the association between stage and emergency presentation with survival using 355,502 patient records, and compared results from this to the published literature. We also analysed the survival of patients missing a recorded stage, who comprise a substantial proportion of patients.

We restricted the literature search to documents explicitly mentioning 'early diagnosis'. This approach gave us insight into what people consider 'early diagnosis' to encompass, however, it excluded studies where explicit mention of 'early diagnosis' was absent, so some data on the association between an indicator and patient survival may have been omitted.

4.7. Conclusion

In this study we identified the different indicators used to measure early diagnosis, and examined the association of each of the indicators with short-term mortality and survival. We recommend several changes to early diagnosis surveillance in England based on our findings: that granular stage information should be used in stage statistics to improve their prognostic value; that patients without a recorded stage should be included in surveillance to minimise bias; and that data on patient's stage and route to diagnosis could be combined to create a composite early diagnosis indicator.

Shorter diagnostic intervals can be a result of late-stage disease and of patients being acutely unwell, and therefore we conclude that the average length of diagnostic interval is not an informative measure for performance management. More work is needed to examine the association between reductions in diagnostic interval length and survival improvements in an area, and to develop informative statistics based on the diagnostic interval for use in surveillance.

Authorship contribution

PM and LW contributed to the study design. PM conducted the literature review, data analysis, and wrote a first draft of the manuscript under the guidance of LW. All authors contributed to interpretation of the study results and to editing the manuscript. All authors approved the final draft for submission.

Conflicts of interest

None.

Acknowledgements

PM and SW received funding from a Cancer Research UK Early Diagnosis Policy Research grant to the Cancer Policy Programme at the London School of Hygiene & Tropical Medicine (award number C7923/A18348), and LW is funded by a Cancer Research UK Population Research Postdoctoral Fellowship (award number C23409/A11415). The authors would like to thank members of the CRUK Scientific Advisory Group for their insights in the development this project. We additionally thank Public Health England for providing the data used in the data analysis, Aurelien Belot for statistical advice, Adrian Turculet for performing the data cleaning and linkage, and members of the LSHTM Cancer Survival Group for many informative discussions.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.canep.2018.07.010.

References

[1] S. Walters, C. Maringe, M.P. Coleman, M.D. Peake, J. Butler, N. Young, S. Bergström, L. Hanna, E. Jakobsen, K. Kölbeck, S. Sundstrøm, G. Engholm, A. Gavin, M.L. Gjerstorff, J. Hatcher, T. Børge Johannesen, K.M. Linklater, C.E. McGahan, J. Steward, E. Tracey, D. Turner, M.A. Richards, B. Rachet, ICBP Module 1 Working Group, Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the United Kingdom: a population-based study, 2004–2007, Thorax 68 (2013) 551–564.

[2] C. Maringe, S. Walters, B. Rachet, J. Butler, T. Fields, P.J. Finan, R. Maxwell, B. Nedrebø, L. Påhlman, A. Sjövall, A. Spigelman, G. Engholm, A. Gavin, M.L. Gjerstorff, J. Hatcher, T. Borge Johannesen, E.J. Morris, C.E. McGahan, E. Tracey, D. Turner, M.A. Richards, M.P. Coleman, ICBP Module 1 Working Group, Stage at diagnosis and colorectal cancer survival in six high-income countries: a population-based study of patients diagnosed during 2000–7, Acta Oncol. 52 (2013) 919–932.

[3] Cancer Care Ontario, Ontario Cancer Plan 2008–2011, 2008.

[4] Cancer Australia, Cancer Australia Strategic Plan 2014–2019, Surry Hills, NSW (2014), p. 40.

[5] Department of Health, Improving Outcomes: A Strategy for Cancer, Department of Health, London, UK, 2011.

[6] Public Health England, Stage at Diagnosis 2012–2014 and One-year Cancer Survival in England, (2016).

[7] Public Health England, Cancer Data Dashboard (2017) (Accessed 17 January 17), https://www.cancerdata.nhs.uk/dashboard#?tab=Overview.

[8] Office for National Statistics, Cancer Survival by Stage at Diagnosis for England (experimental Statistics): Adults Diagnosed 2012, 2013 and 2014 and Followed up to 2015, (2016).

[9] US National Library of Medicine, National Institutes of Health, Pubmed.gov (2017), https://www.ncbi.nlm.nih.gov/pubmed/.

[10] World Health Organisation, International Statistical Classification of Diseases and Related Health Problems – 10th Revision (ICD-10), Geneva, Switzerland (2011).

[11] L. Elliss-Brookes, S. McPhail, A. Ives, M. Greenslade, J. Shelton, S. Hiom, M. Richards, Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets, Br. J. Cancer 107 (8) (2012) 1220–1226.

[12] S. Benitez-Majano, H. Fowler, C. Maringe, C. Di Girolamo, B. Rachet, Deriving stage at diagnosis from multiple population-based sources: colorectal and lung cancer in England, Br. J. Cancer 115 (3) (2016) 391–400.

[13] M.P. Perme, J. Stare, J. Estève, On estimation in relative survival, Biometrics 68 (1) (2012) 113–120.

[14] H. Brenner, O. Gefeller, T. Hakulinen, Period analysis for' up-to-date' cancer survival data: theory, empirical evaluation, computational realisation and applications, Eur. J. Cancer 40 (3) (2004) 326–335.

[15] R. Li, R. Daniel, B. Rachet, How much do tumor stage and treatment explain socioeconomic inequalities in breast cancer survival? Applying causal mediation analysis to population-based data, Eur. J. Epidemiol. 31 (6) (2016) 603–611.

[16] P. Murage, P. Murchie, M. Bachmann, M. Crawford, A. Jones, Impact of travel time and rurality on presentation and outcomes of symptomatic colorectal cancer: a cross-sectional cohort study in primary care, Br. J. Gen. Pract. 67 (660) (2017) e460–e466.

[17] J.M. Ahrensberg, F. Olesen, R.P. Hansen, H. Schroder, P. Vedsted, Childhood cancer and factors related to prolonged diagnostic intervals: a Danish population-based study, Br. J. Cancer 108 (6) (2013) 1280–1287.

[18] S. Gupta, P. Gibson, J.D. Pole, R. Sutradhar, L. Sung, A. Guttmann, Predictors of diagnostic interval and associations with outcome in acute lymphoblastic leukemia, Pediatr. Blood Cancer 62 (6) (2015) 957–963.

[19] M. Laudicella, B. Walsh, E. Burns, P.C. Smith, Cost of care for cancer patients in England: evidence from population-based patient-level data, Br. J. Cancer 114 (11) (2016) 1286–1292.

[20] G. Lurie, L.R. Wilkens, P.J. Thompson, R.K. Matsuno, M.E. Carney, M.T. Goodman, Symptom presentation in invasive ovarian carcinoma by tumor histological type and grade in a multiethnic population: a case analysis, Gynecol. Oncol. 119 (2) (2010) 278–284.

[21] M. Sala, L. Domingo, F. Macià, M. Comas, A. Burón, X. Castells, Does digital mammography suppose an advance in early diagnosis? Trends in performance indicators 6 years after digitalization, Eur. Radiol. 25 (3) (2015) 850–859.

[22] E.A. Tracey, D.M. Roder, D.C. Currow, What factors affect the odds of NSW cancer patients presenting with localised as opposed to more advanced cancer? Cancer Causes Control 23 (2) (2012) 255–262.

[23] J.F. Aitken, M. Elwood, P.D. Baade, P. Youl, D. English, Clinical whole-body skin examination reduces the incidence of thick melanomas, Int. J. Cancer 126 (2) (2010) 450–458.

[24] A. Castanon, R. Landy, P.D. Sasieni, Is cervical screening preventing adenocarcinoma and adenosquamous carcinoma of the cervix? Int. J. Cancer Suppl. 139 (5) (2016) 1040–1045.

[25] D. Ciocan, C. Barbe, F. Aubin, F. Granel-Brocard, D. Lipsker, M. Velten, S. Dalac, F. Truchetet, C. Michel, A. Mitschler, G. Arnoult, A. Buemi, S. Dalle, P. Bernard, A.S. Woronoff, F. Grange, Distinctive features of melanoma and its management in elderly patients: a population-based study in France, JAMA Dermatol. 149 (10) (2013) 1150–1157.

[26] A.J. Chorley, Y. Hirst, C. Vrinten, Cv. Wagner, J. Wardle, J. Waller, Public understanding of the purpose of cancer screening: a population-based survey, J. Med. Screen. 0 (0) (2017) 0969141317699440.

[27] F. Durbec, F. Vitry, F. Granel-Brocard, et al., The role of circumstances of diagnosis and access to dermatological care in early diagnosis of cutaneous melanoma: a population-based study in France, Arch. Dermatol. 146 (3) (2010) 240–246.

[28] A.N. Eskesen, K. Bjoro, E.M. Aandahl, P.D. Line, E. Melum, Low use of surveillance and early diagnosis of hepatocellular carcinoma in Norway-a population-based cohort study, Cancer Epidemiol. 38 (6) (2014) 741–747.

[29] S. Ess, A. Savidan, H. Frick, C. Rageth, G. Vlastos, U. Lutolf, B. Thurlimann, Geographic variation in breast cancer care in Switzerland, Cancer Epidemiol. 34 (2) (2010) 116–121.

[30] R.T. Greenlee, H.L. Howe, County-level poverty and distant stage cancer in the United States, Cancer Causes Control 20 (6) (2009) 989–1000.

[31] J.P. Hreinsson, J.G. Jonasson, E.S. Bjornsson, Bleeding-related symptoms in colorectal cancer: a 4-year nationwide population-based study, Aliment. Pharmacol. Ther. 39 (1) (2014) 77–84.

[32] G. Lurie, P.J. Thompson, K.E. McDuffie, M.E. Carney, M.T. Goodman, Prediagnostic symptoms of ovarian carcinoma: a case-control study, Gynecol. Oncol. 114 (2) (2009) 231–236.

[33] J. Sankaranarayanan, S. Watanabe-Galloway, J. Sun, F. Qiu, E. Boilesen, A.G. Thorson, Rurality and other determinants of early colorectal cancer diagnosis in Nebraska: a 6-year cancer registry study, 1998–2003, J. Rural. Health 25 (4) (2009) 358–365.

[34] World Health Organisation, Guide to Cancer Early Diagnosis, Geneva (2017).

[35] National Cancer Intelligence Network, Colorectal Cancer Survival by Stage, NCIN Data Briefings, England (2009).

[36] Public Health England, Indicator Specification: Proportion of Cancer Admissions Diagnosed for the First Time Via Emergency Presentation, (2015).

[37] Public Health England, Indicator Specification: Cancer Diagnosed at Early Stage: the Proportion of Invasive Malignancies of Breast, Prostate, Colorectal, Lung, Bladder, Kidney, Ovary, Uterus, Non-hodgkin Lymphomas, and Melanomas of Skin, Diagnosed at Stage 1 or 2, (2015).

[38] Independent Cancer Taskforce, Achieving World-Class Cancer Outcomes: A Strategy for England 2015-2020 Independent Cancer Taskforce, London, UK (2015).

[39] G. Lyratzopoulos, C.L. Saunders, G.A. Abel, S. McPhail, R.D. Neal, J. Wardle, G.P. Rubin, The relative length of the patient and the primary care interval in patients with 28 common and rarer cancers, Br. J. Cancer 112 (s1) (2015) S35–S40.

[40] F.M. Walter, G. Rubin, C. Bankhead, H.C. Morris, N. Hall, K. Mills, C. Dobson, R.C. Rintoul, W. Hamilton, J. Emery, Symptoms and other factors associated with time to diagnosis and stage of lung cancer: a prospective cohort study, Br. J. Cancer 112 (s1) (2015) S6–S13.

[41] G. Rubin, C. Gildea, S. Wild, J. Shelton, I. Ablett-Spence, Assessing the impact of an English national initiative for early cancer diagnosis in primary care, Br. J. Cancer 112 (s1) (2015) S57–S64.

[42] M.L. Tørring, M. Frydenberg, R.P. Hansen, F. Olesen, P. Vedsted, Evidence of increasing mortality with longer diagnostic intervals for five common cancers: a cohort study in primary care, Eur. J. Cancer 49 (9) (2013) 2187–2198.

[43] W. Hamilton, S. Stapley, C. Campbell, G. Lyratzopoulos, G. Rubin, R.D. Neal, For which cancers might patients benefit most from expedited symptomatic diagnosis? Construction of a ranking order by a modified Delphi technique, BMC Cancer 15 (2015) 820.

[44] M. Barclay, C. Gildea, J. Poole, L. Hirschowitz, U. Menon, A. Nordin, Factors affecting short-term mortality in women with ovarian, tubal, or primary peritoneal cancer: population-based cohort analysis of english national cancer registration data, Int. J. Gynecol. Cancer 26 (1) (2016) 56–65.

[45] S. McPhail, L. Elliss-Brookes, J. Shelton, A. Ives, M. Greenslade, S. Vernon, E.J. Morris, M. Richards, Emergency presentation of cancer and short-term mortality, Br. J. Cancer 109 (8) (2013) 2027–2034.

[46] R.D. Neal, P. Tharmanathan, B. France, N.U. Din, S. Cotton, J. Fallon-Ferguson, W. Hamilton, A. Hendry, M. Hendry, R. Lewis, U. Macleod, E.D. Mitchell,

M. Pickett, T. Rai, K. Shaw, N. Stuart, M.L. Tørring, C. Wilkinson, B. Williams, N. Williams, J. Emery, Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? Systematic review, Br. J. Cancer 112 (Suppl. 1) (2015) S92–S107.

[47] S.C. Crawford, J.A. Davis, N.A. Siddiqui, L. de Caestecker, C.R. Gillis, D. Hole, G. Penney, The waiting time paradox: population based retrospective study of treatment delay and survival of women with endometrial cancer in Scotland, BMJ 325 (7357) (2002) 196.

[48] R. Neal, Can Earlier Symptomatic Diagnosis Improve Cancer Outcomes in Wales? A Report for Public Health Wales, Bangor University, 2016.

[49] G. Lyratzopoulos, P. Vedsted, H. Singh, Understanding missed opportunities for more timely diagnosis of cancer in symptomatic patients after presentation, Br. J. Cancer 112 (Suppl. 1) (2015) S84–91.

[50] C. Renzi, G. Lyratzopoulos, T. Card, T.P.C. Chu, U. Macleod, B. Rachet, 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 115 (7) (2016) 866–875.

[51] M.E. Barclay, G. Lyratzopoulos, D.C. Greenberg, G.A. Abel, Missing data and chance variation in public reporting of cancer stage at diagnosis: cross-sectional analysis of population-based data in England, Cancer Epidemiol. 52 (Supplement C) (2018) 28–42.

[52] R. Li, L. Abela, J. Moore, L.M. Woods, U. Nur, B. Rachet, C. Allemani, M.P. Coleman, Control of data quality for population-based cancer survival analysis, Cancer Epidemiol. 38 (3) (2014) 314–320.

7.2 Supplementary appendices

Appendix Table 1 Literature search keywords and document inclusion criteria.

Elements*	PubMed search terms**	Google search terms***
cancer	cancer* OR tumor* OR tumour* OR malignan* OR neoplasm*	cancer OR tumor OR tumour OR malignancy OR malignancies OR neoplasm
early diagnosis	"early cancer diagnosis" OR "earlier cancer diagnosis" OR "early diagnosis" OR "earlier diagnosis"	early cancer diagnosis OR earlier cancer diagnosis OR early diagnosis OR earlier diagnosis
population-based	"population-based" OR "population based" OR routine data* OR "routinely collected"	population-based OR population based OR routine data OR routinely collected
Document inclusion criteria		
Criterion (1)	Includes presentation of new statistics or methods for generating statistics based on an explicitly defined indicator of early diagnosis from data sources that are or could be collected routinely	
Criterion (2)	Includes results describing the association between an early diagnosis indicator and patient prognosis (any measure of complications, mortality, survival, life expectancy, or cure).	

* "AND" operator was applied to only return documents containing all three elements

**Title/abstract search

***Different syntax was required to account for the differences in search options and logic implemented between PubMed and Google searches

Appendix Table 2 Number of patients (%) diagnosed with colorectal, non-small cell lung, or ovarian cancer by route to diagnosis, England, 2009-2013.

Route to diagnosis	Colorectal cancer	Non-small cell lung cancer	Ovarian cancer
Death Certificate Only (DCO)	70 (0.0)	163 (0.1)	15 (0.1)
Emergency presentation	37,285 (23.2)	60,632 (35.6)	7,503 (30.7)
GP referral	38,253 (23.8)	35,357 (20.7)	5,291 (21.6)
Inpatient elective*	5,800 (3.6)	2,762 (1.6)	282 (1.2)
Other outpatient†	11,488 (7.2)	18,338 (10.8)	2,597 (10.6)
Screening	15,046 (9.4)	-	-
Two-week wait (TWW)	45,898 (28.6)	45,809 (26.9)	7,485 (30.6)
Unknown‡	3,939 (2.5)	3,352 (2.0)	734 (3.0)
Missing	2,838 (1.8)	4,022 (2.4)	543 (2.2)
Total	160,617 (100.0)	170,435 (100.0)	24,450 (100.0)

* A route to diagnosis commencing with a planned inpatient admission

† An elective route starting with an outpatient appointment

‡ No records of inpatient, outpatient, or screening activity prior to diagnosis

Appendix Table 3 Arithmetic difference between five-year net survival calculated using a 'complete case' approach and a 'missing route is other route' approach for handling missing route to diagnosis data, by age and stage at diagnosis, England 2009-2013.

	All ages	Ages 15-59	Ages 60-79	Ages 80-99
Colorectal cancer				
I	-0.3	0.1	-0.2	-1
II	-0.4	0	-0.4	-0.6
III	-0.4	-0.1	-0.3	-0.6
IV	-0.2	-0.1	-0.2	-0.1
Missing	-0.4	-0.4	-0.4	-0.4
Total	-0.5	-0.2	-0.4	-0.6
Non-small cell lung cancer				
I	-1.1	-0.6	-0.9	-1.3
II	-0.5	-0.7	-0.5	-0.4
III	-0.1	-0.1	-0.1	0
IV	0	0	0	0.1
Missing	0.2	0	0.2	0.1
Total	-0.1	-0.1	0	-0.1
Ovarian cancer				
I	-0.1	0	-0.1	-0.7
II	0	0.3	0	-1.2
III	-0.2	0	-0.2	-0.1
IV	0.2	0.2	-0.1	0.6
Missing	0	0.1	-0.1	0.2
Total	-0.1	0.1	-0.2	-0.1

7.3 Implications for monitoring of trends

Several considerations for monitoring emerge from the literature review and data analysis presented in this Chapter.

As an early diagnosis indicator, the interval from symptoms to diagnosis does not have a simple relationship with survival, with shorter intervals often being due to severe symptoms caused by aggressive or late-stage disease.

Emergency presentation was found to be a proxy for late stage disease, and also associated with lower survival for patients at each stage of diagnosis. It therefore has a broad interpretation as a measure for the success of the health system in preventing adverse effects from cancer prior to diagnosis of it. However, as discussed previously, it is an administrative not a clinical indicator [54], and during 2008-13 the composition of patients diagnosed through the route changed, to include more patients diagnosed following A&E attendances, and fewer via emergency GP referrals [110]. More generally, scenarios where emergency presentations increase without a negative impact on patient's outcomes can be imagined. Patients who struggle to get a GP appointment promptly may decide to attend A&E instead, which could be beneficial for the patients if this expedites diagnosis. Conversely, reductions in emergency presentations could be achieved without immediate benefit to

cancer patients, for example through campaigns to encourage patients to go wait for an appointment with their GP and avoid A&E if at all possible. Such a change might have overall benefits for the health system, but not benefit cancer patients specifically. These nuances make emergency presentation a more complex indicator to interpret compared to stage.

Of those indicators considered, stage is therefore the optimal one for use in monitoring. Whilst it is not completely recorded as emergency presentation is, it is a clinical measure with a clear interpretation and is highly prognostic; not just associated with survival but a causal determinant. Two key considerations for its use stand out from the literature review and data analysis. First, a strategy is needed to reduce bias from the missing data. Second, as every incremental change in stage is prognostic, an analysis of time trends should consider the granular differences in stage alongside, or instead of, any binary “early” vs “late” grouping.

One limitation of this review of indicators is that it considered only their association with patient survival, and not quality of life. For example, the diagnostic interval was not found to have a clear association with survival, and so was not recommended as a subject of routine surveillance. However, from a patient experience perspective, a timely journey from referral with suspected cancer to diagnosis or cancer excluded may significantly improve their experience by reducing time spent waiting for the result. For the purposes of this thesis, however, not all potential indicators were available for analysis, and those with a clear relationship to survival were preferred. With respect to stage, early-stage diagnosis is typically associated with less need for aggressive treatment to achieve cure or remission. This, along with an improved prognosis, generally confers significant improvements in quality of life.

8. Methods for case-mix adjustment

As discussed in the previous Chapter, in order to gain statistics which accurately reflect the performance of health services, appropriate methods are needed to adjust for patient case-mix factors. Numerous statistical techniques exist to perform the adjustment, each enabling comparisons of relative performance which are less biased by the case-mix. In this Chapter different techniques for the adjustment are surveyed and compared, to determine which to employ in the substantive thesis analyses, and meet aim 2 of the thesis: to identify appropriate methods to measure early diagnosis trends with minimal bias.

8.1 Direct standardisation

Direct standardisation is a common approach for age-mix adjustment. It is used by the World Health Organization to adjust for differences in the age structures of countries in international comparisons of mortality [144], and by researchers in international comparisons of cancer survival [145].

In direct age standardisation, for each population compared, a weighted average of results from different age groups is calculated. For example, in a temporal early diagnosis comparison, estimates are generated for each age group in each time period. Weights are then applied to each age group in each of the different periods, and the resulting weighted averages can be interpreted as early diagnosis “as if” the age structure was the same in each time period being compared (the same age structure implied by the weights used).

Direct standardisation can be readily extended to more than one case-mix factor. For two factors one subdivides the groups of the first case-mix factor by levels of the second, so e.g. if standardising on age (categorised into six age groups) and also sex (into two groups), the weighted average would be generated using twelve groups, each defined by one possible combination of age group and sex. Weights can be generated from a pre-specified external reference population, or from the average of all populations being compared in the study (i.e. the average from patients in all time periods in a temporal comparison).

Direct standardisation has the advantage of being simple to calculate, and also easy to interpret. It also has some important limitations. Reasonable numbers of patients must be present in each group in each population compared to return a robust estimate. Estimates can only be generated if there are patients in each group of the population. As the number of groups needed multiplies for each additional case-mix factor which we want to adjust on, it is usually impossible to standardise on more than a few case-mix factors before sparse data becomes an issue, even in relatively large populations.

Direct standardisation also requires all case-mix variables to be categorical. For sex or tumour topography this will be the only logical choice, but for a factor like age, a linear or quadratic relationship may provide a better model for the association with the outcome. If so, the implicit

assumption of a constant association between age and early diagnosis for people of different ages *within* each age group results in residual confounding.

8.2 Indirect standardisation

Indirect standardisation allows adjustment for a larger number of case-mix factors, as there is no need for each combination of case-mix factors to be well represented in each population compared. Instead of weighted averages, the health outcomes in each population are compared to the expected results for patients with those case-mix attributes across all the populations [146]. This contrast can be expressed as the arithmetic difference between observed and expected or, more commonly, a ratio. The technique is prominently applied by the Care Quality Commission (CQC) in England to calculate the standardised mortality ratios used to detect hospital death rates which are unusually high for the mix of patients admitted [147]. Expected values for patients with a specific combination of case-mix attributes can be generated from the average values in the entire population, or estimates from generalised linear models. There is no requirement that every combination of case-mix factors is present in each population, and so no issues arise from sparse data.

In the case of the CQC hospital standardised mortality ratios, for each patient, the "expected" or average number of deaths for someone with their presentation and case-mix characteristics is estimated based on national data. For example, if a hospital admits a 65-year old male who has no pre-existing comorbidities with a myocardial infarction, and nationally 15% of men with these attributes die, then the "expected" deaths ("E") assigned for that man is 0.15. A value of 1 will be assigned to the "observed" deaths if he dies and 0 if not. A ratio for the hospital he was admitted to is calculated based on the sum of all observed deaths ("O") divided by the sum of all expected ("E") from the patients admitted. The formula is as follows [148]:

$$Indirectly\ standardised\ ratio\ (ISR) = \frac{\sum_i O_i}{\sum_i E_i}$$

The 95% confidence interval is calculated by first calculating an error factor ($EF = e^{\frac{1.96}{\sqrt{O_i}}}$), and then calculating the lower limit as $\frac{ISR}{EF}$ and the upper limit as $ISR * EF$ [148].

Similar to direct standardisation, indirect standardisation has the advantage of being straightforward to calculate, and it also returns results with a clear interpretation: each statistic is the percentage difference from the population average, or the reference expected value (when ratios are calculated). Sparse or missing data are not a problem. The major limitation, however, is that as a non-parametric approach no information is returned on total variability, the magnitude of any differences or temporal changes, or the effect of individual case-mix factors. For these to be derived a modelling approach is required.

8.3 Covariable adjustment with generalised linear models

8.3.1 Summary of approach

An alternative to standardisation is for variables representing each level of the exposure of interest (be it hospital, time period, or country) to be put in a generalised linear model (GLM), with the case-mix factors included as covariables. This is more commonly used when a small number of institutions are to be compared (e.g. much less than 100), due to practical challenges of estimating a very large number of parameters within one model. This approach has the advantage that effect estimates and confidence intervals for case-mix factors and the exposures of interest are estimated, and, as with indirect standardisation, there is flexibility in how case-mix factors are specified. For example, they can be included as continuous linear, quadratic, or categorical variables.

The interpretation from this modelling adjustment approach is that the exposure effect estimates from the model are adjusted for the case-mix variables. This adjustment occurs because the estimates returned are based on simultaneous joint optimisation (maximum likelihood estimation) for both the exposure and case mix in the same model, and so each estimate has accounted for the effect of the others.

8.3.2 Challenges for assessing the role of case-mix variables

In addition to interpreting the adjusted estimates from the model, with certain models one can also compare exposure estimates between models which do and do not include the case-mix factors, and then attribute the difference to the effect of case mix [149]. In a linear regression model for a continuous outcome this approach is appropriate. The same approach cannot be done with logistic regression for a binary outcome, however. The problem arises because of the logit link function used in logistic regression:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

where the log odds is modelled, and the effect of a unit increase in variable $x_1$is estimated by the odds ratio e^{β_1}. One attribute of the odds ratios is that an unadjusted odds ratio will be different to the case-mix adjusted odds ratio, *even if the odds ratios are identical in each different group of patients with the same case-mix characteristics*. The weighted average of strata-specific odds ratios will not equal the unadjusted odds ratio, as it would for proportions or continuous quantities [150].

As a consequence of this phenomena (“non-collapsibility”), the effect of age on early diagnosis in a logistic regression model will be different if sex is also included as a covariable in the model, regardless of whether sex is independently associated with both age and early diagnosis, and hence meets the classic criteria for being a confounder. As explained by Hernan *et al* (2011), in the context of confounding and causal inference:

"A quantitative difference between conditional and marginal odds ratios in the absence of confounding is a mathematical oddity (no pun intended), not a reflection of bias. Such difference is irrelevant for the purposes of confounding adjustment because, in the absence of confounding by [a third factor] C, both the conditional and marginal odds ratios are valid. They just happen to be different."[151]

In Chapter 9, multilevel regression models are discussed as a tool for quantifying the total extent of geographic inequalities (total between-area variation), when it is impractical to include a separate parameter for each area in a model. It is relevant to discuss multilevel models in this Chapter to note additionally that estimates of geographic variation derived from a multilevel logistic regression model may be different when case-mix factors are added, even if the case-mix factors are not associated with geographic variation. With multilevel linear regression models, one can estimate the variation explained by directly comparing geographic variation estimates between models which do and do not include case-mix factors [152]. In a logistic model, however, the patient-level variation is solely a function of the mean of the binomial distribution assumed by the model, whilst the between-area variation is estimated on the log-odds scale: the two variances are not directly comparable [148]. The feature that patient-level variation remains fixed also entails that the cluster-level variation will be rescaled if factors which explain individual-level variation are included in the model. The result is that "variation explained" cannot be estimated by directly comparing the models: some inflation of the area-level variance may be expected simply from the addition of the extra patient-level variables. Additional calculations can however be conducted to apply a scale correction factor and calculate variation explained [153].

In summary, covariable adjustment with generalised linear models can be a powerful tool for case-mix adjustment, but there are pitfalls in certain situations. Effect estimates for exposures are returned, as well as estimates of the effect of case-mix factors, along with their p-values and confidence intervals. However, in the case of logistic regression for binary outcomes, variation explained cannot be immediately calculated through direct comparison of effect estimates between models which do and do not also include case-mix factors.

8.4 Conclusion

Direct standardisation is simple to execute and interpret, but sparse data limit its use if too many case-mix factors are included. Indirect standardisation with expected/observed ratios is not limited in this respect. However, estimates of the effects of individual exposures or total variation are not returned by either of these approaches.

Covariable adjustment using GLMs allows for a multitude of case-mix factors and for the exposure of interest to be specified flexibly, and also returns effect estimates (with p-values and confidence intervals) for all the variables included: estimates of exposures, and case-mix factors. There is some need for caution in the case of logistic regression, the typical GLM used for binary outcomes: un-

adjusted and adjusted models cannot be compared directly to understand the effect case mix has on the association between exposure and outcome.

With respect to the conceptual framework outlined in Chapter 6, the number of case-mix variables identified indicates that direct standardisation is likely to be unfeasible but that indirect standardisation or a modelling approach will be feasible. Of these two, only the modelling approach returns estimates (and p-values and confidence intervals) for the individual effects and also has the potential to return estimates of total variability. A modelling approach for case-mix adjustment will therefore be used in the substantive analysis of the thesis, with the noted need for caution with case-mix adjustment and logistic regression in the case of binary outcomes.

9. Approaches to assess geographic early diagnosis inequalities

Geographic inequalities in early diagnosis are well documented [114]. Reducing them has been a priority of recent cancer policy, and proposed as a means to close the survival gap with other countries[3]. The health service changes during 2008-2013 may have had a positive or negative effect. For example, increasingly similar patient pathways due to referral guidelines and waiting times targets may have reduced geographic inequalities. Conversely, differences in uptake of colorectal screening and other initiatives associated with geographic deprivation differences may potentially have led to uneven improvements, increasing inequalities, as has been described in the context of child health and breast and cervical cancer screening interventions in low- and middle-income countries [154, 155]. Monitoring the total extent of geographic inequalities can allow assessment of progress against health service targets and identify areas in need of further investment or support.

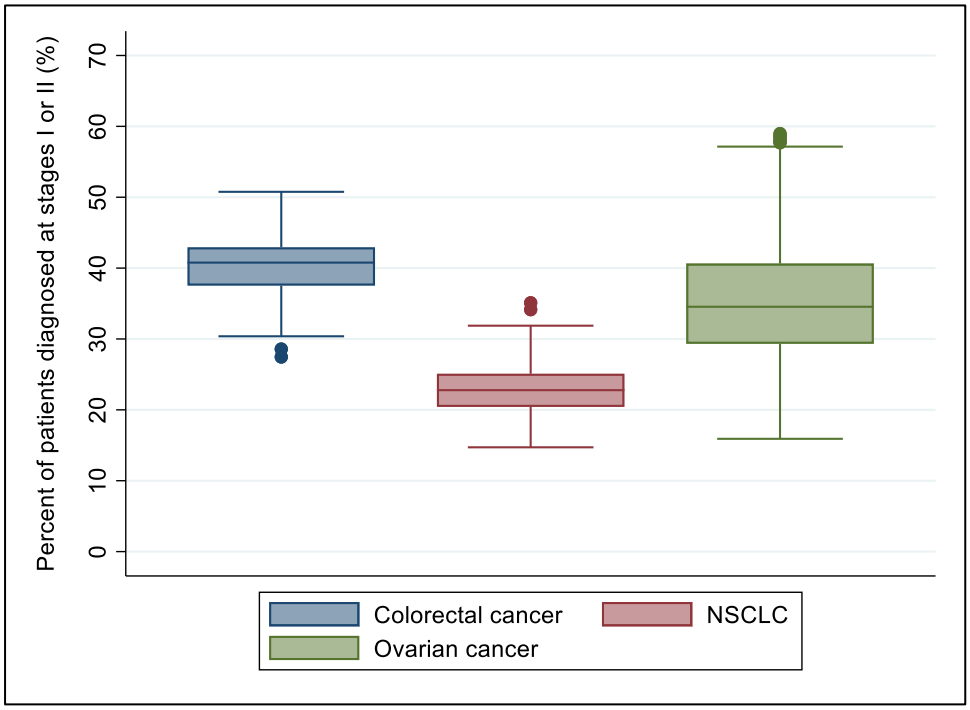
Analysts typically examine inequalities at the CCG level [3, 156], as CCG commissioners have ultimate responsibility for patient outcomes in their area. The contrast between the early diagnosis results for an individual CCG and the national average is part of the formal monitoring of health services conducted through the Public Health Outcomes Framework [157], and CCG-level results are also used for "quality premium" payments [55]. Hence, as well as monitoring the total extent of inequalities to understand their extent and success of interventions to reduce them, individual CCG-level results are of interest in the monitoring of early diagnosis.

This Chapter considers different statistical approaches to analyse geographic inequalities with reference to the literature. The discussion is accompanied by results from data analysis using the techniques applied to colorectal, non-small cell lung, and ovarian cancers during the period 2008-2013. The Chapter concludes with consideration of the most appropriate technique for the substantive analysis in the thesis.

9.1 Descriptive approaches

Variation in CCG-level early diagnosis results can be presented using descriptive techniques. One example is boxplots, which show the median, the 25th and 75th percentiles, and far outliers (Figure 9.1). A boxplot showing (complete case) early-stage diagnosis percentages by CCG indicates substantial variation in ovarian cancer early diagnosis, from 15% to 60%, with less variation for the other two cancers. However, for a robust analysis factors such as random variability and case mix also need to be considered, and other techniques are better suited for this than boxplots.

Figure 9.1 Boxplots showing crude between-CCG variation in early-stage diagnosis in a complete case analysis, England, 2008-2013.



9.2 Funnel plots and the detection of outliers

The benefits of funnel plots in early diagnosis monitoring can be illustrated by considering a simple binary outcome (e.g. early vs late diagnosis), and one CCG. The observed early diagnosis percentage in that CCG (0 – 100 %) can be imagined as a single sample drawn from a wider (hypothetical) patient population in that CCG where there is a certain probability of being diagnosed early i.e. a sample from a population of people getting the same care people in that CCG actually received. With this approach, the observed early diagnosis percentage is subject to sampling error. This conceptualisation is useful, because when patient numbers are small, the early diagnosis statistic is strongly influenced by the random play of chance.

In one extreme example: a CCG might have only one incident ovarian cancer in a given year, and it might happen that the tumour is detected early when the patient is undergoing an ultrasound for another condition. The observed early diagnosis percentage is 100% (1/1). Of course, this result is almost completely uninformative as to its performance and the outcomes future patients in that CCG can expect. However, if the data is considered a sample (n=1) from a hypothetical population, statistical techniques for expressing uncertainty around estimates from limited samples can then be used. These return confidence intervals and p-values to quantify the precision of the estimate from the sample. This approach is (implicitly) routinely used in surveillance of population-level data, including in cancer survival and early diagnosis comparisons, where p-values and confidence intervals are estimated [5, 18].

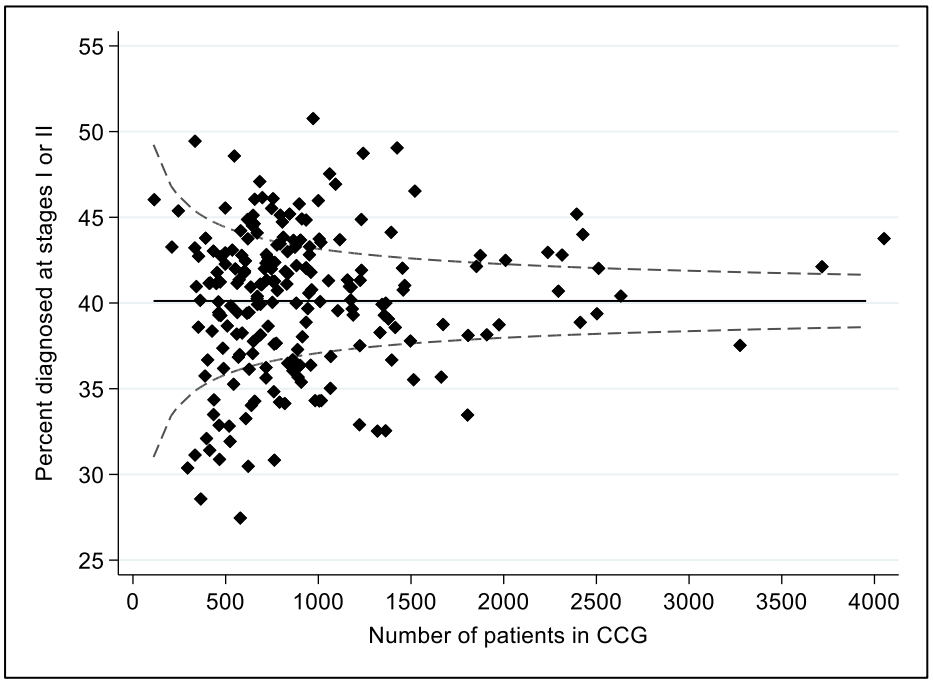
Statistical sampling theory is also useful for interpreting performance against targets and the spread of results from different CCGs. Observed variation in the early diagnosis percentage is heavily influenced by the sample size used to calculate the percentage. For example, consider the following situation:

- Nationally, the overall probability of early diagnosis is 50%;
- In one selected CCG, the probability of early diagnosis is also 50% (in line with the national average).

In an analysis project to identify CCGs where over 70% or less than 30% of patients were diagnosed early for further investigation, the likelihood that this CCG is chosen depends on the early diagnosis percentage sample size. If 10 patients from that CCG are randomly sampled, then the chance the *observed* early diagnosis percentage is outside the 30%-70% range is high: over a third (34%) of the time the observed value will fall outside this range, even though the actual probability any patient is diagnosed early is 50%, simply due to chance. However, if a sample of 100 patients is taken, there is a very small (<1%) chance the early diagnosis percentage for this CCG will be outside the 30%-70% range. So, even if a CCG has typical performance, with a small sample an extreme result can be returned just by the random play of chance.

More generally, some effect of sampling error is expected even with large samples. For example, if 2,000 patients are analysed from the CCG with 50% early diagnosis probability, the observed percentage would still frequently be as low as 48% or as high as 52% just due to sampling error (based on the 95% confidence interval for a proportion, n=2000, p=0.5). Given that even for common cancers such as colorectal cancer no more than 150 patients are diagnosed in an average CCG in a year [158], random chance must be considered as a dominant factor in observed variation in early diagnosis during 2008-2013, even before other considerations such as case-mix differences.

Figure 9.2 Funnel plot for between-CCG variation in colorectal cancer diagnosis at stages I or II (complete case analysis), with the scatterplot of CCG-level results overlaid on 95% control limits, England, 2008-2013



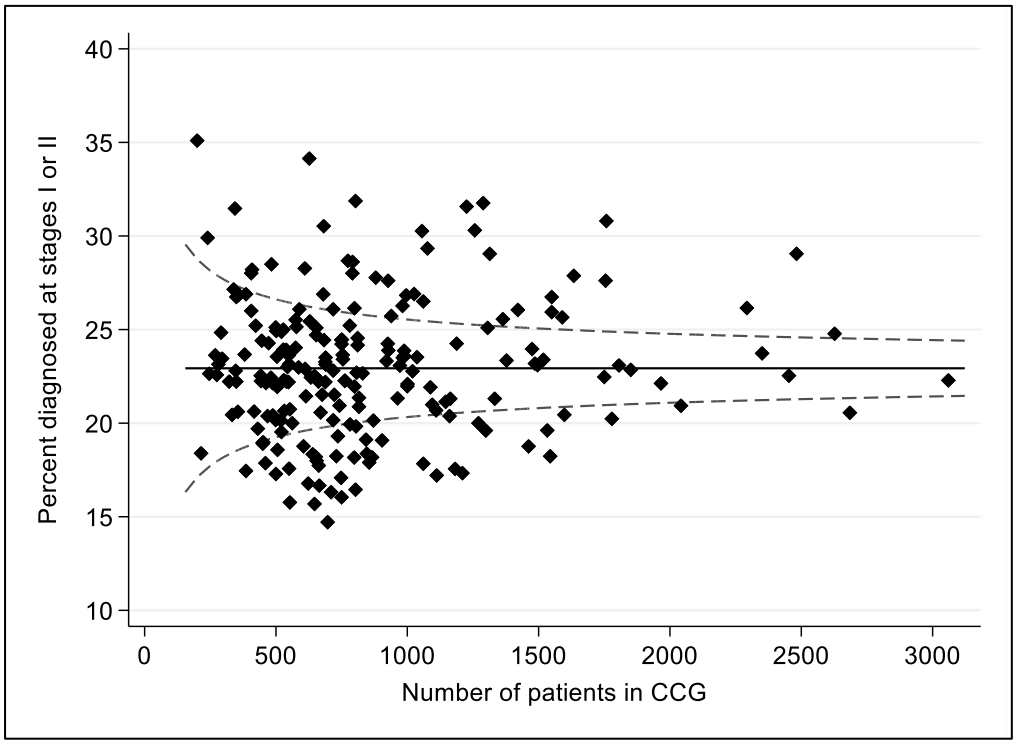
Funnel plots take account of differences expected due to sampling variability. With these, a reference value is chosen (usually the national or whole-population average), and control limits around it are defined for each possible sample size, n, with limits chosen to reflect the uncertainty due to sampling error; being wider where n is small and narrower where n is large [147]. The 95% confidence intervals around the reference value are typically chosen as the control limits.

To construct the funnel plot, results for each CCG are plotted, on the y-axis, against the sample size on which the result is based, on the x-axis, with control limits shown in the background (illustrated in Figure 9.2). Results fall outside the control limits only if they are significantly different to the reference at the 5% level (i.e. a hypothesis test that the early diagnosis percentage in the CCG is equal to the national average returns $p<0.05$). Informally, if results fall inside the control limits then there is no evidence the CCG has different early diagnosis to the national average. If results fall outside the funnel limits, there is evidence the CCG has different early diagnosis from the national average (at the $p=0.05$ level) [147].

Funnel plot results can be analysed in different ways. One approach is to designate all results outside the control limits as anomalous or noteworthy. There can be problems with this approach, however. The 95% control limits are calculated such that, when a CCG has an early diagnosis probability equal to the average, there is a 5% chance its results will fall outside the control limits by random chance. In a single comparison of one CCG and the national average, the result falling outside the control limits might be considered enough evidence for the analyst to conclude early diagnosis in the CCG is different to the average (given $p<0.05$). On a funnel plot with results for all 211 CCGs in 2013, even if

all actually had probability of early diagnosis equal to the average, ~10 (5%) would be expected to fall outside the limits simply by chance, as this is the definition of the 5% control limit. Hence, a single result outside the limits does not by itself provide firm evidence a CCG is different to the average, and several results outside the limit do not constitute evidence for excess variation. Furthermore, even when there is evidence that a CCG result is different from the average, there may not evidence that the difference is large.

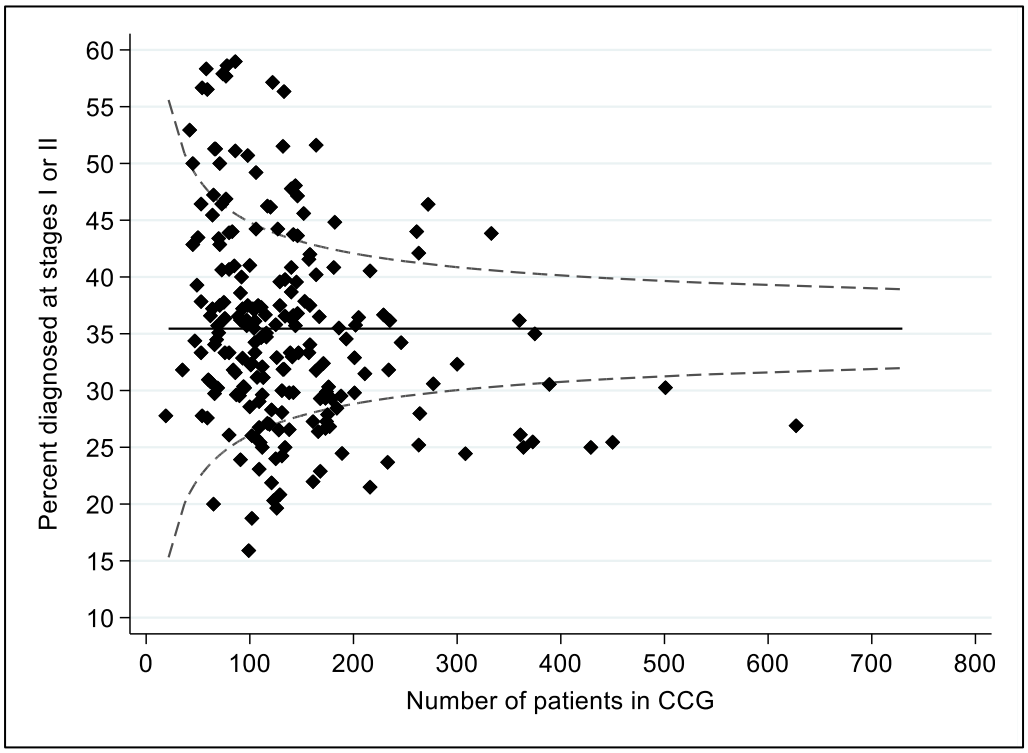
Figure 9.3 Funnel plot for between-CCG variation in NSCLC diagnosis at stages I or II (complete case analysis), with the scatterplot of CCG-level results overlaid on 95% control limits, England, 2008-2013



Some other approaches with funnel plots include extending the control limits (e.g. 99.8% instead of 95%); identifying only very extreme outliers; or identifying persistent outliers across time periods or consistent outliers across indicators [159]. For extreme values, the plots can help analysts visually identify results which are not only (statistically) significantly different from the average but which also stand out against the general patterns of variability between areas.

In the complete case funnel plot analysis of the early diagnosis (stage I or II) percentage in 2008-2013 for each of colorectal cancer, non-small lung cancer, and ovarian cancer, presented in this Chapter, the results show that for each cancer there is large excess variability around the national average, but also no stand-out extreme values outside the general funnel shapes (Figures 9.2, 9.3, 9.4).

Figure 9.4 Funnel plot for between-CCG variation in ovarian cancer diagnosis at stages I or II (complete case analysis), with the scatterplot of CCG-level results overlaid on 95% control limits, England, 2008-2013.



In addition to sampling error, case-mix differences may also explain some of the observed between-CCG variation in the early diagnosis. These differences are adjusted for in the funnel plots of national clinical audit reports, where case-mix adjusted trust-level outcomes are typically presented instead of raw results [160].

For directly standardised risk statistics, control limits based on the standard error of the estimate can be used instead of n on the x-axis. Use of these preserves the property that results fall outside the control limits only if there is evidence the result is different from the average. The sample size n does not meet this criteria so cannot be used, because there is no 1:1 relationship between the sample size and the sampling error (which defines the control limit values) for directly standardised percentages. One alternative quantity, which is more easily interpretable than the standard error, is n*= SE^2/((target)*(1-target)), where SE is the standard error of the risk-adjusted estimate [161]. n* is interpretable as “approximately the sample size”.

For indirectly standardised estimates such as observed/expected events ratios, the same approach can be taken: control limits can be generated based on the confidence interval around the reference value (i.e. 1, in the case of the ratio).

9.3 Modelling total variability and estimating reliability

Another approach for assessing geographic differences is to estimate the total variation with a statistical model. The control limits on a funnel plot are themselves a representation of the simplest model, in which the only differences observed are due to sampling error. If this model is found to fit the data poorly (e.g. if substantially more than 5% of values fall outside the 95% control limits, as is the case in Figures 9.2-9.4), an adjustment can be made to expand the control limits to account for this over-dispersion, and ensure roughly 5% of values fall outside the control limits [159]. This is done by adding another parameter to the model to account for this additional (unwanted) variation. The remaining variation can be interpreted as the "true" geographic variation; that not due to random chance, but instead genuine organisational performance and case-mix differences.

The decomposition of variation into that due to chance and that due to underlying variability can also be achieved with multilevel generalised linear models (GLMs). These estimate total between-CCG variability, by assuming the spread of CCG effects are normally distributed around the national average. The multilevel modelling approach reduces the number of parameters needed in the model, as there is just one parameter for the overall variation instead of one for each CCG. This also means no direct information on any one CCG is returned, though post estimations can be done to extract these [162]. It is also possible to extend this approach to include parameters for geographic variation in different time periods, and use a Wald or equivalent likelihood to test whether total inequalities changed.

Total variance estimates from multilevel models can also be used to calculate reliability statistics. These estimate the proportion of variability due to between-organisation (i.e. CCG of residence) differences as opposed to within-organisation differences (i.e. between patients with different characteristics within a CCG). The reliability is calculated as $\mu/(\mu + \sigma)$, where μ is the between-organisation variability (estimated in the random effects model) and σ is the patient-level variability [44]. Reliability takes a value from 0 to 1, where 0 indicates all observed variability is due to patient-level differences, and 1 that it is all due to between-organisation variability. A higher reliability means that underlying between-CCG differences play a more significant role in the geographic variation observed, whereas lower values imply these are less important. Some heuristics recommended for England are that an average reliability of 0.7 should be required for statistics which are reported publicly, and 0.85 should be required when the statistics are used for pay for performance schemes [44, 163].

In one analysis of variation in early-stage diagnosis by CCG, it was estimated that the median reliability of early diagnosis percentages calculated using data from 10 common cancers in 2013 was only 0.66, below the threshold for public reporting [44]. With respect to the target set in 2015 for 60% of tumours to be diagnosed at stages I or II, Barclay *et al* also estimated that the positive predictive value of a CCG returning an early diagnosis percentage over 60% was only 53% (i.e. at a reliability of 0.66, nearly half would be falsely misclassified as meeting the target when their true percentage was <60%). When still using combined data from 10 common cancers, it was estimated that 3 years of

data was needed for CCG-level reliability to be consistently over 0.7, and 8 years be needed for reliability of over 0.9. It was thus recommended that CCG-level statistics should not be included in pay for performance schemes.

For this thesis, CCG-level reliability statistics were estimated separately for each of three cancers studied in the 2008-2013 period using a random-intercept logistic regression model, with patient-level variance estimated as $\pi^2/3$ following the latent-normal response interpretation [153]. With reference to the heuristics for public reporting used by Barclay *et al*, acceptable reliability for public reporting (but not pay for performance schemes) was achieved when using six years of combined data for each of colorectal and NSCLC, but not for ovarian cancer (Table 9.1). These results indicate that, even with six years of data, caution is needed when assessing performance of a CCG in early diagnosis by simply comparing its observed early diagnosis against a fixed target.

Table 9.1 Average reliability of early-stage percentage statistics for 209 CCGs by cancer type, with between-CCG variation estimated from a random intercept logistic regression model, England, 2008-2013.

Malignancy	Average reliability
Colorectal cancer	0.72
Non-small cell lung cancer	0.81
Ovarian cancer	0.56

9.4 Conclusion

CCG-level early diagnosis statistics are heavily influenced by sampling variation. An initial analysis of reliability statistics returned evidence that comparing observed early diagnosis in a CCG against a target is a reasonably accurate way of measuring its performance against that target, for only two of the three cancers I examined, before any consideration of case mix.

Funnel-plot analyses can account for sampling variation, but results from these can also be challenging to interpret. A first un-adjusted analysis of geographic variation for the three cancers I considered returned evidence for excess variability around the national average, and also no signs that results for any CCG were extreme outliers. Case mix and missing data were not addressed in these initial analyses, and may explain (or mask) some of the geographic early diagnosis differences.

Despite sparse data limiting the inferences that can be made about the performance of individual CCGs, modelling approaches can be used to estimate the overall between-CCG variation. Overall

variation can also be compared between time periods, facilitating evaluation of whether efforts to reduce geographic inequalities are successful.

10. Missing data and modelling

In this Chapter the strategy for the substantive analysis is detailed, prior to presentation of the results in Chapter 11. Several decisions for this analysis have been outlined previously: case-mix factors to be adjusted for include tumour subsite and morphology, age, sex, and comorbidity; and stage will be the outcome.

Missing data is one key challenge for the substantive analysis. It is a concern for the outcome, stage, but also for the case-mix factors topographical subsite and morphology. Multiple imputation is introduced in this Chapter as an approach to handle missing data, and its application to the substantive analysis is outlined. The specification of the imputation models in R, and the multilevel analysis models for substantive analysis in Stata, are then described in detail.

10.1 Multiple imputation to minimise bias from missing data

During 2008-2013 stage at diagnosis data was missing for 19.9% of NSCLC patients, 37.4% of colorectal cancer patients, and 40.7% of ovarian cancer patients in England. Higher percentages were missing stage in the earlier years 34-62% of patients with these three cancers lacked a record of stage in 2008-2009, compared to only 9-19% in 2012-2013. Any systematic differences between the stages of patients with and without stage recorded could potentially introduce substantial bias into estimates of trends. However, the technique of multiple imputation can be used to reduce bias from the missing data. It achieves this by making more plausible assumptions about the similarity of patients with and without stage recorded, compared to a complete case analysis.

10.1.1 Different mechanisms for missing data

The mechanisms by which data is missing can be categorised into one of the following: missing completely at random (MCAR), missing at random (MAR), or missing not at random (MNAR) [164]. Which one of these is true determines which statistical approach (if any) will provide un-biased and accurate results. If data are MCAR, the probability data are missing is completely unrelated to the value they take, and a complete-case analysis will return unbiased results. If data are MAR, data are missing randomly conditional on values of other variables which are recorded. In that case, a complete-case analysis will be biased in certain situations, but multiple imputation can be used to minimise bias. The final mechanism, MNAR, is when the probability of data for a variable being missing is associated with the underlying values taken. In this case, an un-biased analysis is typically impossible, but sensitivity analyses may be done to evaluate how inferences from the analysis change under different scenarios.

10.1.2 Overview of multiple imputation

Multiple imputation is a popular approach for handling missing data when the MAR assumption is reasonable [165]. In brief, a model fitted using data from patients with non-missing values for the

variable of interest (the 'imputation model') is used to produce a distribution of plausible values for each item of missing data, based on the association between that variable and all the others which may be predictive of it. A number of different complete datasets are then created from all the non-missing data items and, for the items of missing data, from random draws from these distributions of plausible values. A model to answer the research question ('analysis model') is then fitted on each of these complete datasets. A final set of parameter estimates are created by averaging results from the same analysis model run on the different datasets.

Two components determine the total error about the final set of model parameter estimates and confidence intervals: the estimated within-model error (i.e. the error from estimating parameters using a generalised linear model), and the between-model error (based on the variability of results between different imputation datasets). Compared to simpler alternatives, multiple imputation provides unbiased parameter estimates, better estimates of total variability, and best statistical power - if the MAR assumption holds and the imputation model is correctly specified [166].

Multiple imputation is useful for retrieving information from 'auxiliary' variables, which predict the values of missing data and whether data are missing, but which are not planned for use in the analysis model. As an example, in a study on income and cancer survival, some patient records might have missing values for income, with data on income more likely to be missing for patients with higher incomes. If other data associated with income is available (e.g. educational level, occupation, housing type, area of residence), these may predict income so accurately that, for patients who are the same with respect to these, there are no systematic differences in income regardless of whether it was a recorded or not. The auxiliary variables cannot be meaningfully included alongside income as co-variables to be adjusted for in the analysis model. However, they can be included in the imputation model to predict income where missing, they will greatly minimise bias from missing data in the analysis.

In this scenario – where stage is the outcome, and there exist auxiliary variables which can be used to recover information about it when it is missing – multiple imputation can provide more accurate estimates. Multiple imputation can also be useful to recover information and minimise bias when stage is an analysis covariable, for example in analyses of net survival by stage [167]. However, in situations where additional auxiliary variables are not available to improve the imputation (i.e. where the only variables in the imputation model are those already included in the analysis model) and stage is the outcome (and only analysis variable with missing data), a complete case analysis will return very similar estimates to multiple imputation, and may be preferable [168].

For the assumption of MAR to be plausible, the imputation model should include all variables associated with the value of the missing data items, or the likelihood of data being missing, even if the association is weak [166]. The rationale is as follows: if a variable not associated with values of missing data or data being itself is erroneously included there is minimal loss of accuracy, yet inclusion of a variable which is associated will reduce variability in the values imputed, reducing the

between-model component of overall variability, resulting in smaller confidence intervals and p-values.

10.1.3 Plausibility of the MAR assumption for English stage data

Di Girolamo *et al* have described variables in population-based datasets which are strongly associated with missing information on patient's stage [115]. For the multiple imputation models used in the thesis substantive analysis, all the variables identified in that study and additional variables required for the analysis model were included:

Diagnosis time: year and quarter-year

Location: Cancer Network area, and CCG area of residence

Demographics: age, sex, IMD income quintile, Charlson comorbidity score

Tumour: topography and morphology (where known; set to a 'unspecified' or 'non-specific' category in <12% of records)

Route to diagnosis: Emergency presentation, GP referral, 2-week wait, outpatient, inpatient, unknown, screening (route used for colorectal cancer only)

Treatment: Major surgical treatment within [-30, 90] days of diagnosis; treatment admission method (defined as 'elective' versus 'emergency')

Survival: time to censoring, vital status at censoring.

For the MAR assumption to hold, where stage is missing it must be missing randomly *conditional* on equality in the above variables. Equivalently, there must not exist a mechanism by which stage is systematically different between any two sub-groups of patients sharing identical values of the above variables.

A 'reduction to absurdity' approach can help assess the MAR assumption: assume it is not true and derive a contradiction. For MAR not to hold with the above auxiliary variables, some unaccounted for factor must be negating the very strong casual association between stage and survival time [97, 169]. This could be:

i) A factor shortening the survival of a group of the patients with earlier stage disease, so they have equal survival time to patients with later stage disease;
ii) A factor extending the survival of a group of the patients diagnosed at later stages, so they have equal survival time to patients with early-stage disease.

Regarding a mechanism for scenario i), one possibility is that some patients may refuse treatment or fail to attend appointments, leading to shorter survival time and failure to ascertain or record stage. Inclusion of information on major surgical treatment in the imputation models would be expected to reduce bias from this mechanism, however, where present.

Another mechanism for i) or ii) is differences in tumour grade. Direct information on tumour grade is too poorly recorded to be usefully included in the imputation model to minimise any bias from this mechanism for NSCLC and ovarian cancer (at 17% and 32% completeness respectively), but grade is recorded for 67% of the colorectal cancer patients who are missing stage, and could reduce bias for these patients. Inclusion of morphology in the imputation model will also reduce any bias from this mechanism, to the extent that morphology is a proxy for grade, and for those registrations where morphology is completely recorded.

A third potential mechanism is differences in treatment not captured by the major treatment variable (which only measures major surgical treatment), such as chemotherapy and radiotherapy provision. Any differences in access or uptake of these between patients with similar prognostic characteristics could introduce bias.

Bias could also be introduced if the association between the above predictor variables and stage changes during the study period, and the imputation model fails to account for this. For example, if stage-specific survival improves over time and the model does not account for this, longer survival in the later years may be erroneously attributed to earlier stage rather than, for example, the effect of improved treatment. Concern about this type of error could be addressed by inclusion of an additional interaction term between diagnosis period and survival time, which would allow for different relationships between survival and stage at different times within 2008-2013.

Considering these potential mechanisms for bias, and the reasons stage is missing outlined in Chapter 5, most of the key prognostic variables associated with the different reasons for stage data being missing are included. The inclusion of survival time, alongside these, should substantially reduce – if not entirely eliminate - bias.

10.1.4 Number of complete datasets created

Bodner *et al* recommend a rule of thumb that the 'the number of imputations should be similar to the percentage of cases that are incomplete', based on a simulation study which showed this ensures high precision in the model parameter estimates, p-values and confidence intervals [170]. White *et al* additionally recommend this rule because, compared to even higher numbers of imputations, it provides sufficient statistical efficiency, very small loss of power to detect an association of interest, and reproducibility of results should the entire imputation process be repeated with different draws from estimated distributions for missing values [166].

10.1.5 How much missing data is too much to impute?

The maximum percentage of missing data it is acceptable to handle using multiple imputation has been discussed in the literature, though without a definite conclusion.

One simulation study evaluated changes in the root mean square error (RMSE) when the extent of missing data varied from 0% to 80% [171]. The RMSE is a composite function of the variance and the bias of the estimator. Lower results indicate lower bias and variance, and therefore better performance. In one example where data were MAR, the RMSE did increase with the proportion of

missing data when using MI, but the increase was slight, and the effect estimates remained unbiased at very high levels of missing data. In this case the variance around estimates increased in proportion to the missing data fraction, but this did not lead to false inferences: rather it accurately represented the uncertainty introduced from imputing a large percentage of data. By contrast, in a simulation where the MAR assumption did not hold, the RMSE increased dramatically with the missing data percentage, and bias was introduced into the estimation.

White *et al* corroborate these findings: they state that in theory any fraction can be validly imputed, provided the imputation is done correctly and the MAR assumption is correct [166], but any imperfections will have proportionately larger impact when large fractions are imputed. They conclude:

"It would seem wise to take special care if more than 30-50 percent are to be imputed."

For this thesis, the substantive analysis was done using multiple imputation even for cancers for which >30% of the stage data is missing. As this choice increased the potential for residual bias, sensitivity analyses were planned to further evaluate the robustness of the results, and caution was exercised not to over-interpret the results.

10.1.6 Ensuring the imputation model is at least as complex as the analysis model

One further consideration is the need for the multiple imputation model to account for the multilevel structure of the data, in order to analyse geographic inequalities. Analysis model parameter estimates are biased if the imputation model does not include them, and so all parameters in the analysis model should be included in the imputation model. This rule applies equally to cluster-level variables in a hierarchical model (i.e. CCG of residence) as to patient-level variables: use of an imputation model that does not account for between-CCG differences could lead to underestimates of geographic inequalities in the substantive analysis, as stage will be (implicitly) imputed assuming there are no geographic inequalities.

An imputation model may account for geographic inequalities either by including a parameter for each geographic territory (which is likely to be computationally challenging in this case, due to the large number of parameters required) or by using a hierarchical model. Traditionally, multiple imputation models could not account for hierarchical data, however the recently developed R package *jomo* solves the issue of incompatibility between (non-hierarchical) imputation and (hierarchical) analysis models using a joint modelling approach [172]. *Jomo* was used to create the imputed datasets for each cancer for the thesis analyses. The use of *Jomo* allows the probability of early diagnosis to vary between CCGs, but otherwise assumes the association between the different patient-level variables and stage at diagnosis is the same in each CCG, mirroring the assumptions of the analysis model.

10.2 Recoding of variables for the imputation and analysis models

There were two key considerations for recoding of variables in the imputation and analysis models:

- **Parsimony:** including all variables that are predictive of stage or missing data, without unnecessary or collinear parameters which could increase error around parameter estimates or stop the imputation or analysis models converging.
- **Appropriate handling of non-specific morphology and topography values:** As described in Chapter 5, the topography and morphology variables have "non-specific" values, which should be considered as missing and not included as another category of the case-mix variable, as they are likely to be downstream from late diagnosis.

The rationales for the different choices for certain variables are described below.

10.1.1 Topography

Colorectal cancer

Colorectal cancer topography codes were regrouped as follows:

Table 10.1 Re-grouping of colorectal cancer topographical sites.

Topography	Missing stage %	Stage 1/2 % *	Total	New group
Right colon (C18.0-18.4)	38.1%	39.9%	63,261	Colon
Left colon (C18.5-18.7)	39.8%	41.5%	52,620	Colon
Overlapping lesion of colon (C18.8)	38.0%	37.4%	505	Colon
Colon, NOS (C18.9)	58.0%	20.5%	10,766	Colon
Rectosigmoid junction (C19)	41.7%	36.0%	13,439	Rectum
Rectum (C20, C21.8)	35.7%	43.6%	55,920	Rectum

*Of patients with a recorded stage

To ensure parsimony with variables the colorectal topographies were regrouped into 'colon' and 'rectum' sites; the same groupings used in the Office for National Statistics Cancer Survival statistics [56]. The 'Colon, NOS' topography code may be a proxy for late-stage disease, and is uninformative as regards the actual topography of a tumour within the subsite. Under the groupings chosen there is no need to recode it to missing however, as a C18.9 code indicates a tumour which originates somewhere in the colon (in the site C18), and therefore if fixed to missing could only be validly imputed back to 'Colon'.

Non-small cell lung cancer (NSCLC)

The codes 'Lung, NOS' were re-coded to missing, as were overlapping lesions, as both indicate a tumour originating in one of the other specific sites. The exact lobe the tumour originated within the

lung was not associated with material differences in missing stage or early-stage diagnosis probability, so patients with tumours originating in the lower, middle, and upper lobes of the lung were grouped together into one category. Patients with tumours originating in the main bronchus were substantially more likely to have missing stage information than those with a tumour originating in one of the lobes. They were also far less likely to be diagnosed at stage 1 or 2. These patients were therefore kept as a separate group.

Table 10.2 Re-grouping of NSCLC topographical sites.

Topography	Missing stage	Stage 1/2 *	Total	New group
Main Bronchus (C34.0)	18.8%	7.8%	8,953	Main Bronchus
Upper lobe lung (C34.1)	13.3%	28.0%	74,173	Lobe
Middle lobe lung (C34.2)	15.2%	28.4%	5,873	Lobe
Lower lobe lung (C34.3)	14.9%	29.1%	39,207	Lobe
Overlapping lesion (C34.8)	17.8%	27.1%	800	[missing]
Lung, NOS (C34.9)	34.7%	10.0%	51,042	[missing]

*Of patients with a recorded stage

Ovarian cancer

Ovarian cancer topography did not need re-coding, as it uses only two codes (C56 – Ovary, C57 – fallopian tube).

Table 10.3 Grouping of ovarian cancer topographical sites.

Topography	Missing stage	Stage 1/2 *	Total	New group
Ovary (C56)	40.8%	33.3%	28,181	[no change]
Fallopian tube (C57)	34.7%	46.4%	895	[no change]

*Of patients with a record stage

10.2.2 Morphology

Similar to the ‘Not otherwise specified’ topographies for NSCLC, it was necessary to recode the ‘non-specific’ morphologies to missing and then impute them, as these categories will frequently be downstream from late-stage diagnosis.

Colorectal cancer

As a pragmatic move to reduce the number of categories, neuroendocrine neoplasms were grouped together with other specific morphologies and mesenchymal tumours into a new 'Non-carcinoma' group, reducing the number of morphology categories to two (Table 10.4). These new groupings have a simple clinical interpretation, and also relatively well discriminate the patients with higher early-stage diagnosis and less missing data (carcinomas). Though the mesenchymal tumours have different missing stage probability non-carcinoma tumours, these comprise a tiny proportion of all tumours and so are unlikely to have a material impact on results.

Table 10.4 Grouping of colorectal cancer morphologies.

Morphology	Missing stage	Stage 1/2 *	Total	New group
Carcinomas	36.1%	40.8%	156,743	Carcinoma
Neuroendocrine neoplasm	58.5%	41.6%	2,783	Non-carcinoma
Mesenchymal tumours	82.1%	50.0%	56	Non-carcinoma
Other specific morphologies	43.1%	63.0%	16,175	Non-carcinoma
Non-specific morphologies	56.9%	12.8%	20,754	[missing]

*Of patients with a recorded stage

Non-small cell lung cancer (NSCLC)

Similarly, for NSCLC, the non-specific morphologies were re-coded to missing. The same simple division of morphologies into carcinomas and non-carcinomas as was implemented for colorectal cancer was used. For NSCLC this division discriminates well the morphologies with similarly high early-stage diagnosis and lower missing stage from the others.

Table 10.5 Regrouping of NSCLC morphologies.

Morphology	Missing stage	Stage 1/2 *	Total	New group
Adenocarcinoma	13.0%	24.8%	49,958	Carcinoma
Squamous cell carcinoma	11.5%	28.8%	39,512	Carcinoma
Large cell carcinoma	15.4%	27.4%	2,303	Carcinoma
Other specific morphologies	24.1%	17.9%	70,253	Non-carcinoma
Non-specific morphology	43.2%	22.5%	18,198	[missing]

*Of patients with a recorded stage

Ovarian cancer

Ovarian cancer morphology is frequently divided into two groups: type I, which develop from benign extraovarian lesions which subsequently become malignant but are of lower grade, and type II, which arise from intraepithelial carcinomas of the fallopian tube and are higher grade [173].

Table 10.6 Regrouping of ovarian cancer morphologies

Morphology	Missing stage	Stage 1/2 *	Total	New group
Serous carcinoma	31.00%	20.70%	12,004	Type II epithelial
Mucinous	33.40%	82.00%	1,904	Type I epithelial
Endometrioid	24.40%	79.80%	1,978	Type I epithelial
Clear cell	27.50%	67.30%	1,468	Type I epithelial
Other epithelial-stromal	34.00%	39.90%	1,125	Type II epithelial
Unclassified epithelial	54.20%	13.40%	7,937	Type II epithelial
Germ cell	53.90%	74.70%	479	Non-epithelial
Sex cord-stromal	43.50%	82.70%	306	Non-epithelial
Misc and unspecified	80.30%	19.20%	1,875	[missing]**

*Of patients with known stage

** Except for the 86 patients whose initial morphology subgroup was ‘other specific non-epithelial’, who were mapped to the Non-epithelial category.

Matz *et al* (2017) used a division of 6 distinct morphological types for ovarian cancer to examine morphology differences globally, derived based on clinical consultation and previous work by the National Cancer Intelligence Network [174]. Ovarian cancer morphologies were re-categorised to those same groupings for the thesis substantive analysis (Table 10.6). This division well-discriminates the epithelial morphologies with low missing stage and high early-stage diagnosis.

Additional steps were then taken to simplify the categories and reduce the number of imputation and analysis model parameters. Patients with the morphology subgroup ‘other specific non-epithelial’ (86 patients total in 2008-2013) were separated from their parent category ‘Misc and unspecified’, leaving only the ‘Unspecified’ morphologies which were then recoded to true missing.

The 86 patients with 'other specific not epithelial' morphology subgroup were then grouped together with patients with Germ Cell and Sex cord Stromal morphologies into a 'Non-epithelial' category. Though the Germ Cell and Sex cord Stromal tumours are not biologically similar, and have very different mean age of diagnosis, they were grouped together pragmatically to reduce the number of parameters. It was considered that the low numbers of patients with these morphologies and a similar stage 1 or 2 percentage meant any choice of grouping would have little material impact on the case-mix adjustment.

10.2.3 Recoding of other variables

Cancer registry and Cancer Network

Both Cancer Registry area (8 in England), and Cancer Network area (28 in England) were considered for inclusion in the imputation model, as both could plausibly be associated with either stage missingness or early diagnosis probability. However, there is some redundancy in including both, and geographic variation is already accounted for by the inclusion of random effects for patient's CCG of residence. It was decided to remove Cancer Networks from the imputation model, but to retain Cancer Registry areas. Historic Cancer Registry boundaries are not exactly coterminous with CCG boundaries, but they may have significant influence missing stage probability during 2008-2013, as historically the different registries took special interest in different cancers [114].

Charlson comorbidity index

Very small numbers of patients had over three major comorbidities recorded in HES: less than 5% of NSCLC patients, and less than 3% of colorectal or ovarian cancer. The number of comorbidities was therefore grouped into 0, 1, 2, and 3+.

Major treatment and major treatment admission method

The frequency of major treatment was inspected for each cancer. It varied substantially between malignancies, with 61% of all colorectal cancer patients receiving treatment compared to only 12% of NSCLC patients (Table 10.7). The major treatment records also contain information on the admission method. This can take one of 18 values from four overarching groups:

- **Elective:** Waiting list, booked, or planned
- **Emergency admission:** A&E or dental casualty, GP request, bed bureau (from residential care settings), consultant clinic, via mental health crisis resolution team, A&E department of another provider, transfer from another provider in an emergency
- **Maternity admission:** admitted ante- or post-partum
- **Other admission:** birth of a baby, baby born outside the provider, transfer of admitted patient (non emergency)

The vast majority of major treatments (80-98%) were elective admissions (Table 10.7). The rest were overwhelmingly emergency admission, with trivial numbers of major cancer treatments having

admission methods of ‘Maternity admission’ or ‘Other admission’. Treatment admission method was therefore grouped into ‘Elective’ and ‘Non-elective’ for the imputation model.

Table 10.7 Receipt of major cancer treatment and treatment admission method, by cancer type, England, 2008-2013.

Major treatment admissions	Colorectal	NSCLC	Ovary
Not treated	76,415	179,005	12,865
Treated (elective admission)	95,933	23,720	14,968
Treated (non-elective admission)	24,163	482	1,386
Total	**196,511**	**203,207**	**29,219**
Patients receiving major treatment	61%	12%	56%
Percentage of those with an elective admission type	80%	98%	92%

There was a very strong correlation between the ‘emergency presentation’ route to diagnosis and the ‘emergency’ major treatment method for colorectal cancer and ovarian cancer. However, there were over 2,000 patients with each cancer who had an emergency route to diagnosis and non-emergency treatment admission code, indicating that the patient did either not require admission at first emergency attendance, or was admitted and stabilised before being discharged then readmitted or transferred. It was therefore determined that the admission method variable can provide valuable extra information which is predictive of survival, and should be included in the imputation model.

10.3 Implementation of the imputation and analysis models

10.3.1 Imputation models

Missing stage was imputed on the basis of auxiliary patient information using joint modelling with the R package *jomo* (subcommand *jomo1rancat*) [172], treating stage as a categorical variable with no ordering and accounting for the multi-level structure of the data.

The models impute stage where missing separately for each malignancy, by assuming it is missing randomly conditional on quarter year of diagnosis, cancer registry area, CCG, age, sex, patient’s Indices of Multiple deprivation (IMD) income quintile, Charlson comorbidity score [27], tumour topography, tumour morphology, route to diagnosis, receipt of major treatment (yes/no), major treatment admission method (elective/non-elective), time from diagnosis to censoring in days, and vital status at censoring.

Information on survival time was included in the imputation model by the inclusion of i) the Nelson-Aalen cumulative hazard estimate and ii) a binary indicator for whether censored or died. Three

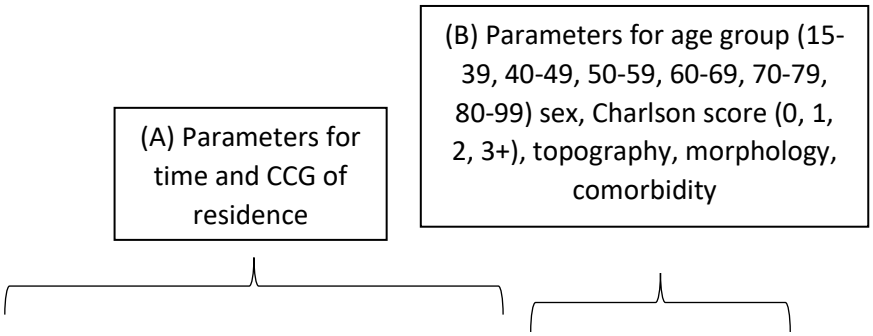
random effects were included, one for CCG in each of the periods 2008-09, 2010-11, and 2012-13 (i.e. each period assigned an independent random intercept, each estimated from the subset of patients diagnosed in that time period). Cancer registry area, deprivation, Charlson score (0, 1, 2, 3+), age group (15-39, 40-49, 50-59, 60-69, 70-79, 80-99), topography, morphology, route to diagnosis, and quarter year of diagnosis (Jan-Mar 2008, Apr-Jun 2008, …, Jul-Sep 2013, Oct-Dec 2013) were all included as categorical variables.

The imputation model thus allowed for differences in the stage distribution between periods as small as three months, while geographic differences were estimated using longer two-year periods, to allow for more precise estimates of geographic inequalities in each period (and hence comparison between them).

The number of imputation datasets created was 39, 20, and 41 respectively for colorectal cancer, NSCLC, and ovarian cancer, in line with the proportion of missing data. These numbers were chosen to reduce the computational burden whilst achieving minimal power reduction compared to using n=100 datasets (an estimated reduction of <1% for all cancers [175]). Parameter estimates (percentages of patients at different stages and regression model parameters) were produced using each dataset and combined using Rubin's rules [165].

10.3.2 Analysis models

The analysis models were specified as follows, where i denotes a patient, j their CCG of residence; y_{ij} whether a patient was diagnosed at stage I or II, $\boldsymbol{X_{ij}}$ the vector of patient and tumour covariables, and μ_{jk} the variance associated with CCG effects on early diagnosis in period k:



$$logit\left(\Pr\left(y_{ij}=1\middle|\boldsymbol{X_{ij}},\boldsymbol{U_{jk}}\right)\right)=\left(\beta_1+\mu_{j1}\right)+T_2(\beta_2+\mu_{j2})+T_3(\beta_3+\mu_{j3})+\beta_4x_{4i}+\cdots+\beta_nx_{ni}$$

$$\mu_{jk}\sim N(0,\varphi_k^2)$$

$$T_2=\begin{cases}1\ if\ diagnosis\ date\ is\ in\ 2010\ or\ 2011;\\ 0\ \ otherwise\end{cases}$$

$$T_3=\begin{cases}1\ if\ diagnosis\ date\ is\ in\ 2012\ or\ 2013;\\ 0\ otherwise\end{cases}$$

The models were fitted with just the (A) parameters (no case-mix adjustment) then with both the (A) and (B) parameters (case-mix adjustment). The (B) parameters were specified in the analysis model as categorical variables, identical to their specification in the imputation model.

The odds ratios exp(β_2) and exp(β_3) provide information on changes in the probability of diagnosis at stage I/II at the national level. The total estimated between-CCG variance at each time period ($\mu_{j1},\mu_{j2},\mu_{j3}$), and the derived odds ratios for 2.5th, 25th, 75th, and 97.5th percentiles of CCG effects, provide information on geographic inequalities.

The command *meqrlogit* was used to fit the regression models in Stata, as this command typically has fewer problems with convergence compared to the alternative melogit when estimated random effects are small [176]. The model estimates distinct CCG effects for each period as separate levels in the model (with no overlap in records used to estimate effects at different levels), in an analogous specification to the multilevel heterogeneity models used in longitudinal studies which allow variability in growth curves between boys and girls [162].

11. Temporal and geographic changes in stage at diagnosis in England during 2008-2013: A population-based study of colorectal, lung, and ovarian cancers

To meet the third thesis aim of evaluating changes in early cancer diagnosis, and changes in geographic inequalities, during 2008-2013, an empirical data analysis was conducted. The analysis uses stage as the outcome, and the optimal methods for case-mix adjustment, geographic comparison, and missing data handling described in Chapters 8-11. The project was published as a research report in the journal *Cancer Epidemiology* in 2020 and is presented in full in this Chapter, along with the online web appendices. The Chapter concludes with further consideration of the strengths and limitations of the analysis, a discussion of the consistency of the results reported with other evidence, and an outline of remaining questions and outstanding work for the thesis.

11.1 Published manuscript

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	270040	**Title**	Mr
First Name(s)	Patrick		
Surname/Family Name	Muller		
Thesis Title	Statistical approaches for monitoring early cancer diagnosis in England		
Primary Supervisor	Dr Laura Woods		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?	Cancer Epidemiology		
When was the work published?	August 2020		
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	**Yes**	Was the work subject to academic peer review?	**Yes**

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	
Please list the paper's authors in the intended authorship order:	
Stage of publication	Choose an item.

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I was the lead author of this study. I planned the study, conducted the analysis, and prepared the draft of the paper. Co-authors provided feedback and input on the study design and the manuscript draft.

SECTION E

Student Signature	P Muller
Date	10/05/2021

Supervisor Signature	L Woods
Date	10/05/2021

Cancer Epidemiology 67 (2020) 101743

Contents lists available at ScienceDirect

Cancer Epidemiology

journal homepage: www.elsevier.com/locate/canep

Temporal and geographic changes in stage at diagnosis in England during 2008–2013: A population-based study of colorectal, lung and ovarian cancers

Patrick Muller*, Laura Woods, Sarah Walters

Cancer Survival Group, Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, Keppel Street, London, WC1E 7HT, United Kingdom

ARTICLE INFO

Keywords:
Cancer
Stage at diagnosis
Early diagnosis of cancer
Time trends
Temporal changes
Geographic inequalities
Population surveillance
Case-mix adjustment
Missing data
Multiple imputation

ABSTRACT

Background: Increasing diagnosis of cancer when the disease is still at early stages is a priority of cancer policy internationally. In England, reducing geographical inequalities in early diagnosis is also a key objective. Stage at diagnosis is not recorded for many patients, which may bias assessments of progress. We evaluate temporal and geographical changes in stage at diagnosis during 2008–2013 for colorectal, non-small cell lung, and ovarian cancers, using multiple imputation to minimise bias from missing data.

Methods: Population-based data from cancer registrations, routes to diagnosis, secondary care, and clinical audits were individually linked. Patient characteristics and recorded stage were summarised. Stage was imputed where missing using auxiliary information (including patient's survival time). Logistic regression was used to estimate temporal and geographical changes in early diagnosis adjusted for case mix using a multilevel model.

Results: We analysed 196,511 colorectal, 180,048 non-small cell lung, and 29,076 ovarian cancer patients. We estimate that there were very large increases in the percentage of patients diagnosed at stages I or II between 2008-09 and 2012-13: from 32% to 44% for colorectal cancer, 19% to 25% for non-small cell lung cancer, and 28% to 31% for ovarian cancer. Geographical inequalities reduced for colorectal and ovarian cancer.

Interpretation: Multiple imputation is an optimal approach to reduce bias from missing data, but residual bias may be present in these estimates. Increases in early-stage diagnosis coincided with increased diagnosis through the "two week wait" pathway and colorectal screening. Epidemiological analyses from 2013 are needed to evaluate continued progress.

1. Introduction

Diagnosis of cancer when the disease is still at an early stage is associated with markedly improved survival prospects [1,2]. Increasing the proportion of patients diagnosed at early stages (often defined as stages I or II) is a focus of cancer policy in the UK and internationally [3–9].

In England, increasing early diagnosis has been identified as one means to reduce the survival gap with other affluent countries [10]. Numerous early diagnosis targets and interventions have been initiated. In 2000 a target was introduced that no patient should have more than a two-week wait (TWW) to see a cancer specialist following general practitioner (GP) referral with possible cancer symptoms [7]. From 2007 that target was extended to include patients referred from a hospital or through screening [11]. In 2005 national guidance for GPs on referring patients with possible cancer symptoms to specialists was published [12]. Faecal Occult Blood Test (FOBT) screening for colorectal cancer was rolled out nationally during 2006–2009 [13], and from 2011 the 'Be Clear on Cancer' campaign has raised awareness of the symptoms from common and rarer cancers, and encouraged people to report them to their GP [14]. In 2015 it became national policy that by 2020 62% of staged cancers should be diagnosed at stages I or II; that the proportion of cancers staged should increase; and that inequalities between the local healthcare commissioners (Clinical Commissioning Groups - CCGs) should decrease [3].

To monitor progress against these targets, from 2016 Public Health England have produced a public-facing website of cancer statistics, the *CancerData* dashboard [15]. The dashboard presents the percentage of cancers with the disease stage recorded, and the percentage of those diagnosed at stages I or II, nationally and for each CCG, for each year from 2012. This "stages I or II" percentage is needed to monitor progress against the target set in 2015. However, it may be biased if used

* Corresponding author.
E-mail addresses: patrick.muller@lshtm.ac.uk (P. Muller), laura.woods@lshtm.ac.uk (L. Woods), sarah.walters@lshtm.ac.uk (S. Walters).

https://doi.org/10.1016/j.canep.2020.101743
Received 14 November 2019; Received in revised form 4 May 2020; Accepted 5 May 2020
Available online 11 June 2020
1877-7821/

for analyses of changes in stage in the whole population, as it excludes patients whose stage was not ascertained or not collected centrally. Nationally, stage recording increased dramatically from 2008 but still only covered 71% of patients in 2013 [15]. The patients without stage recorded have poorer outcomes than patients with stage recorded, suggesting a less favourable underlying stage distribution [16,17].

In addition to missing data, a consideration when interpreting stage trends is the extent to which observed changes are due to changes in health services, for example the introduction of a screening programme, or patient case mix, such as a decrease in the incidence of hard-to-detect tumours. Case-mix differences have been found to confound CCG rankings of early-stage diagnosis [18], and may also influence temporal comparisons.

In this study we analyse temporal and geographic differences in stage at diagnosis during 2008–2013, using statistical techniques to account for missing data and case mix differences. We analyse three malignancies commonly diagnosed late: colorectal cancer, non-small cell lung cancer (NSCLC), and ovarian cancer. We evaluate whether the number of patients diagnosed at stages I or II increased; whether geographic inequalities increased or decreased; and whether observed changes are associated with case-mix. Multiple imputation is employed to minimise bias from missing stage data [19–21].

2. Materials and methods

2.1. Data

Data on cancer registrations were obtained from the Office for National Statistics (ONS) for adults aged 15–99 years, diagnosed with colorectal cancer, NSCLC or ovarian cancer in England from 1 January 2008 to 31 December 2013 (ICD-10 codes C18-20, C21.8; C33-34; and C56-C57.7 [22]). Data on patient's vital status was complete up to 31 December 2014. Data were additionally linked to the national bowel and lung cancer audit datasets [23,24], the Routes to Diagnosis (RtD) dataset [25], and Hospital Episodes Statistics (HES) records using patient's NHS number and postcode. The audit datasets were used to gain additional information on stage [26], whilst RtD records provided information about patient's interactions with the National Health Service (NHS) before diagnosis. The HES records provided information on receipt of major surgical treatment following diagnosis (based on OPCS Classification for Interventions and Procedures version 4 codes; full list in the supplementary appendices) and Charlson Comorbidity Index (CCI - derived from HES records from 6 to 60 months prior to cancer diagnosis) [27].

Clinical Commissioning Group (CCG) areas were used to examine geographical inequalities. These territories were chosen as they have been responsible for commissioning cancer services from 2013, following the dissolution of the Primary Care Trusts which were previously responsible for cancer treatment. Differences in the proportion of patients diagnosed at an early stage were compared between three time periods: 2008−09, 2010−11, and 2012−13. Two-year periods were chosen to ensure each had sufficient numbers of patients for a robust comparison of geographic inequalities.

2.2. Descriptive analysis

Temporal changes in the distribution of stage at diagnosis (I, II, III, IV, or missing) were evaluated. For those with a recorded stage, the percentage of patients diagnosed at stage I or II was tabulated by patient characteristics. The association between each characteristic and missing stage was assessed.

2.3. Multiple imputation

Multiple imputation was conducted to estimate patients' stage of disease at diagnosis where unknown [28]. Imputation models including auxiliary patient information were fitted with the R package *jomo* [29], which accounts for the multi-level structure of the data (patients clustered within CCGs). It was assumed that stage was missing randomly conditional on variables strongly associated with either stage (I to IV) [16,30–32], or with recording of stage [17]: quarter year of diagnosis, cancer registry area, CCG, age, sex, patients' Indices of Multiple Deprivation (IMD) income quintile, Charlson Comorbidity Index, tumour topography, tumour morphology, route to diagnosis, receipt of major surgical treatment (yes/no), treatment admission method (elective/non-elective), time from diagnosis to censoring, and vital status at censoring. Cancer registry area of diagnosis was included as well as CCG, as historically the regional registries recorded stage at different levels of completeness for different tumours.

Tumour morphology and topography included categories which were uninformative as to the actual values ("non-specific", "miscellaneous and unspecified"). These values were re-coded to true missing and imputed using *jomo* alongside missing stage.

The number of imputation datasets created was equal to the percentage of missing data for each cancer: 39, 20, and 41 respectively for colorectal cancer, NSCLC, and ovarian cancer. These numbers are sufficient to achieve a $< 1\%$ power reduction compared to using $n = 100$ datasets [33]. Parameter estimates (percentages of patients at different stages, and regression model parameters) were produced using each dataset and combined using Rubin's rules [34]. Full details of the imputation are provided in the supplementary appendices.

2.4. Regression modelling

The change in the odds of diagnosis at stages I or II between the two-year time periods was estimated using multilevel logistic regression models. Parameters to estimate the between-CCG variation in diagnosis at stages I or II in each time period were fitted and compared using Wald tests. These were also used to estimate odds ratios for CCGs at the 2.5th, 25th, 75th, and 97.5th percentiles, to illustrate the differences in early diagnosis odds between average CCGs and those with highest and lowest percentages of patients diagnosed early. The models were fitted in STATA using *meqrlogit* [35].

The first set of models included only time period and CCG as explanatory variables. A second set of case-mix adjusted variables were fitted including these variables along with age, sex, CCI, tumour topography, and tumour morphology. Further details on the model specification are provided in the supplementary appendices.

3. Results

We analysed cancer registrations of 196,511 colorectal, 180,048 NSCLC, and 29,076 ovarian cancer patients diagnosed during the period 2008–2013 (Table 1). On average in each time period and in each CCG there were 313, 287, and 46 new diagnoses of colorectal cancer, NSCLC, and ovarian cancer respectively (Appendix Table 1).

3.1. Descriptive analyses

Amongst patients with stage recorded, the percentage diagnosed at stages I or II increased dramatically over time for colorectal cancer (from 31.0% in 2008-09 to 45.0% in 2012−13) and NSCLC (from 20.0% to 25.6%), whilst for ovarian cancer it remained similar (changing from 33.3% to 32.9%).

For colorectal cancer, lower deprivation, higher Charlson score, diagnosis following screening or GP referral, and non-carcinoma disease morphology were all associated with early-stage diagnosis (Table 1). For NSCLC, factors associated with early diagnosis included female sex, higher Charlson score, diagnosis following referral from a GP or outpatient service, tumour origin in the lobe as opposed to main bronchus, and carcinoma morphology. For ovarian cancer, early diagnosis was associated with younger age, lower Charlson score, type I

Table 1
Numbers of patients and percentage diagnosed at stages I or II by age, sex, diagnosis period, deprivation, comorbidity, tumour topography, tumour morphology, route to diagnosis, cancer registry, major treatment and admission method.

	Colorectal cancer				NSCLC				Ovarian cancer		
	Count	Stage I/II (%)*	Missing stage (%)		Count	Stage I/II (%)*	Missing stage (%)		Count	Stage I/II (%)*	Missing stage (%)
Total											
All patients	196,511	40.5	39.2		180,048	23.1	20.1		29,076	33.8	40.7
Age											
15 – 39	3,458	37.5	41.8		865	31.4	29.6		1,368	64.3	38.7
40 – 49	7,423	33.3	35.3		4,158	18.5	17.9		2,551	50.3	34.5
50 – 59	20,763	36.0	33.5		17,099	19.5	14.8		4,905	42.2	33.1
60 – 69	50,801	41.7	36.9		46,325	23.2	15.6		7,743	30.2	36.6
70 – 79	60,785	42.3	37.4		61,184	24.6	18.7		7,181	25.6	41.4
80 – 99	53,281	40.2	46.0		50,417	22.6	27.5		5,328	21.5	55.9
Sex											
Male	110,042	40.4	38.0		100,176	22.0	19.7				
Female	86,469	40.5	40.7		79,872	24.5	20.5				
Diagnosis period											
2008 – 2009	63,972	31.0	63.0		57,382	20.0	34.2		9,618	33.3	57.1
2010 – 2011	66,113	39.8	39.0		60,103	22.5	18.1		9,863	35.4	45.7
2012 – 2013	66,426	45.0	16.4		62,563	25.6	8.9		9,595	32.9	19.0
Deprivation quintile											
1 (Least deprived)	42,040	41.8	39.3		25,083	22.7	21.2		6,109	33.5	39.2
2	43,913	40.7	38.6		32,024	23.1	20.3		6,465	31.8	39.8
3	41,033	40.4	39.1		35,995	22.1	20.3		6,228	34.3	42.1
4	36,972	39.5	39.8		39,872	22.9	19.9		5,560	33.4	41.6
5 (Most deprived)	32,553	39.4	39.2		47,074	24.2	19.2		4,714	36.5	40.8
Charlson comorbidity Index											
0	156,968	39.9	38.4		123,622	20.7	19.2		24,896	34.5	39.6
1	18,596	43.1	41.8		28,211	29.0	21.1		2,244	29.6	44.5
2	11,347	41.3	41.1		13,873	28.1	21.9		1,145	32.3	47.0
3+	9,600	43.4	44.8		14,342	28.1	23.8		791	21.7	52.2
Topography											
Colon	127,152	39.4	40.5	Main Bronchus	8,953	7.8	18.8	Ovary	28,181	33.3	40.8
Rectum	69,359	42.2	36.9	Lobe	119,253	28.4	13.9	Fallopian tube	895	46.4	34.7
Missing				Missing	51,842	10.4	34.4				
Morphology											
Carcinoma	156,743	40.8	36.1	Carcinoma	91,659	26.6	12.4	Type I epithelial	5,350	77.1	28.4
Non-carcinoma	19,014	60.6	45.5	Non-carcinoma	70,198	17.9	24.1	Type II epithelial	21,066	19.7	39.9
Missing	20,754	12.8	56.9	Missing	18,191	22.5	43.2	Non-epithelial	873	76.7	52.3
								Missing	1,787	17.2	80.5
Route to diagnosis											
GP referral	45,760	43.7	37.6		37,907	30.5	18.1		6,319	42.5	39.5
Two-week wait	54,249	41.3	29.7		46,647	24.2	8.5		8,600	38.1	28.5
Emergency presentation	44,631	26.4	42.8		64,131	12.7	26.8		9,008	15.4	49.6
Inpatient elective	7,337	42.6	38.9		2,724	14.6	19.9		366	38.5	43.2
Other outpatient	14,027	44.9	41.5		19,510	36.9	18.8		3,172	45.8	40.8
Screening	16,557	59.6	32.6								
Unknown	13,950	32.7	75.5		9,129	24.8	42.7		1,611	33.4	58.9
Registry											
North & York	27,139	41.4	38.4		31,102	25.4	19.7		3,501	33.7	26.7
Trent	20,082	38.4	52.5		19,096	24.1	22.9		2,806	44.9	55.4
East Anglia	23,718	39.9	35.1		19,076	20.1	16.2		3,560	25.9	52.6
Thames	36,843	40.6	50.5		32,892	21.2	24.5		5,768	32.8	43.8
Oxford	10,360	42.7	48.3		7,987	25.1	24.9		1,657	40.8	44.2
South & West	31,065	41.5	29.1		23,551	21.1	18.4		4,636	29.1	28.7
West Midlands	21,324	39.7	30.5		18,412	21.9	16.2		3,290	33.6	38.2
Northwest & Mersey	25,980	39.7	33.0		27,932	26.0	18.5		3,858	39.1	41.8
Major treatment											
Yes	120,096	50.0	35.2		23,804	80.6	11.7		16,353	46.4	29.5
No	76,415	22.7	45.5		156,244	13.2	21.3		12,723	8.4	55.0
Treatment admission method**											
Elective	95,933	53.9	34.3		23,334	80.9	11.7		14,967	46.2	29.1
Non-elective	24,163	33.4	38.7		470	66.8	16.2		1,386	48.1	34.3

* Of patients with a recorded stage.
** Of patients who received major treatment.

epithelial or non-epithelial disease, and diagnosis following referral from a GP or outpatient service. For all cancers, receipt of major treatment and elective admission for treatment were strongly associated with early diagnosis.

Overall, stage was not recorded for 39.2%, 20.1%, and 40.7% of colorectal, NSCLC, and ovarian cancer patients respectively. The percentage of patients missing stage decreased dramatically over time for all three cancers, from 34.2-63.0% in 2008-09 to 8.9-19.0% in

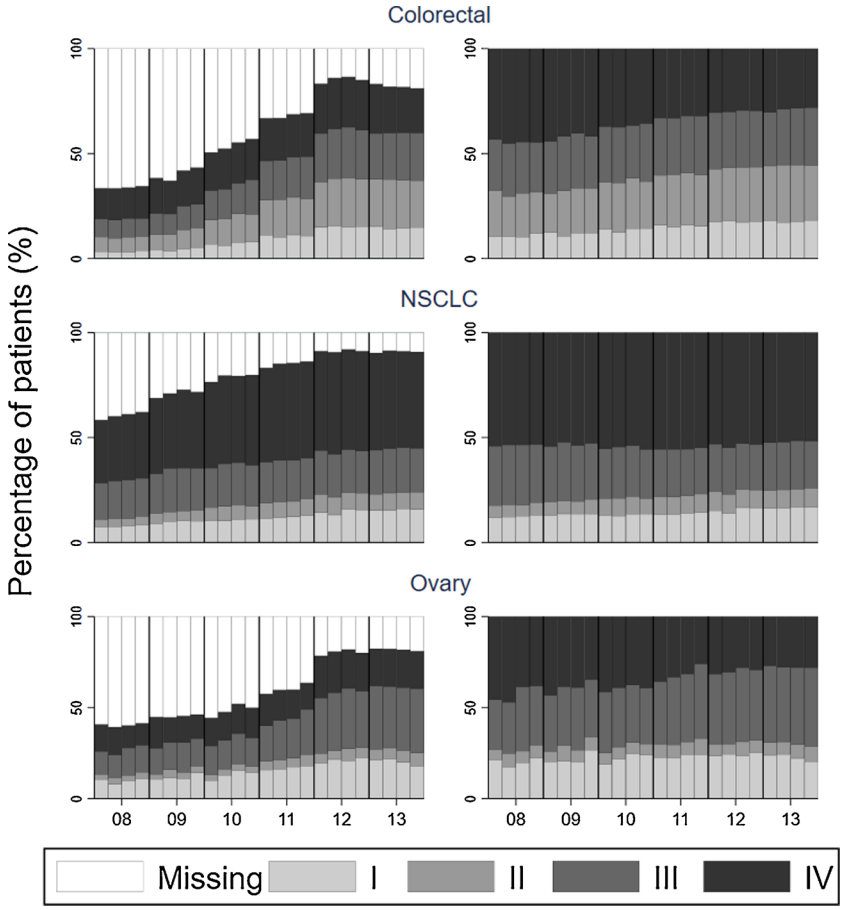


Fig. 1. Distribution of stage at diagnosis in England: comparison of crude results (left) and distribution based on multiple imputation (right) by quarter-year of diagnosis, 2008-2013.

2012 – 13. As stage recording improved, the prognostic characteristics of patients without a recorded stage became less favourable: emergency presentation and pre-existing comorbidities became more common (Appendix Table 2).

Lack of a recorded stage was more common amongst patients who were very young or old compared to the rest of the cohort. Pre-diagnosis it was associated with pre-existing comorbidities and the emergency diagnosis. Post-diagnosis it was associated with a lower probability of receiving major treatment, and with non-elective (unplanned) admission. Lack of recorded stage was also associated with absence of records on tumour topography and morphology.

3.2. Temporal changes

Multiple imputation-based estimates indicate large increases in the percentage of colorectal cancer patients diagnosed at stages I or II nationally, from 32% in 2008-09 to 44% in 2012-13 (compared to from 31% to 45% in the complete case analysis; Fig. 1). For NSCLC the stages I or II percentage increased from 19% to 25% (compared to 20% to 26%). For ovarian cancer it rose from 28% to 31% (compared to remaining at 33%). These estimates also provide evidence of a stage shift from IV to III for colorectal and ovarian cancers: the percentage of stage III tumours amongst all stages III or IV rose from 37% to 48% for colorectal cancer and from 44% to 59% for ovarian cancer (Table 2).

For NSCLC, early diagnosis is estimated to have increased between 2008 – 09 and 2012 – 13 in both models with and without case mix adjustment (Table 3, Fig. 2). However, it was a smaller increase in the case-mix adjusted model (OR: 1.26 (95% CI: 1.20, 1.32) compared to 1.40 (95% CI: 1.33, 1.47), Table 3). During 2008–2013 the case mix for NSCLC shifted towards carcinomas (59.3% by 2012 – 13 compared to 49.5% in 2008 – 09, Appendix Table 3)), tumours originating in a lobe, and patients with pre-existing comorbidities; all characteristics associated with earlier diagnosis.

By contrast, for ovarian cancer early diagnosis is estimated to have increased in both models, but it was a greater increase in the model in

Table 2
Distribution of stage from multiple imputation estimates, by period and cancer.

	2008 – 2009	2010 – 2011	2012 – 2013
		Colorectal cancer	
Patients diagnosed (N)	**63,972**	**66,113**	**66,426**
Stage I/II % (95% CI)	31.9 (31.4, 32.5)	38.5 (38.1, 39.0)	43.8 (43.4, 44.3)
Stage I % (95% CI)	11.3 (10.8, 11.8)	14.6 (14.3, 15.0)	17.7 (17.3, 18.0)
Stage II % (95% CI)	20.6 (20.0, 21.2)	23.9 (23.5, 24.3)	26.2 (25.8, 26.5)
Stage III % (95% CI)	24.9 (24.4, 25.5)	26.9 (26.4, 27.3)	26.8 (26.4, 27.1)
Stage IV % (95% CI)	43.1 (42.5, 43.8)	34.6 (34.1, 35.1)	29.4 (29.0, 29.8)
		NSCLC	
Patients diagnosed (N)	**57,382**	**60,103**	**62,563**
Stage I/II % (95% CI)	18.9 (18.5, 19.3)	21.7 (21.3, 22.0)	24.8 (24.5, 25.1)
Stage I % (95% CI)	12.9 (12.6, 13.2)	13.5 (13.2, 13.7)	16.1 (15.8, 16.4)
Stage II % (95% CI)	6.0 (5.8, 6.3)	8.2 (8.0, 8.5)	8.7 (8.5, 8.9)
Stage III % (95% CI)	27.7 (27.2, 28.2)	23.3 (22.9, 23.6)	22.5 (22.1, 22.8)
Stage IV % (95% CI)	53.4 (52.9, 53.9)	55.1 (54.6, 55.5)	52.7 (52.3, 53.1)
		Ovarian Cancer	
Patients diagnosed (N)	9,618	9,863	9,595
Stage I/II % (95% CI)	27.9 (26.7, 29.1)	29.7 (28.7, 30.8)	30.5 (29.5, 31.4)
Stage I % (95% CI)	21.1 (20.0, 22.2)	22.8 (21.9, 23.8)	23.4 (22.5, 24.3)
Stage II % (95% CI)	6.8 (6.2, 7.5)	6.9 (6.2, 7.6)	7.0 (6.5, 7.6)
Stage III % (95% CI)	31.7 (30.2, 33.2)	34.8 (33.6, 36.0)	40.7 (39.6, 41.8)
Stage IV % (95% CI)	40.4 (39.0, 41.9)	35.5 (34.3, 36.6)	28.8 (27.8, 29.8)

Table 3
Multi-level logistic regression results: Odds ratios (OR) for change in geographic inequalities and in probability of diagnosis at stages I or II during 2008-13.

	No case mix adjustment			Case mix adjustment done***		
	2008 – 09	2010 – 11	2012 – 13	2008 – 09	2010 – 11	2012 – 13
			Colorectal cancer			
OR for difference between time periods (95% CI)	1.00	1.34 (1.26, 1.42)	1.70 (1.61, 1.80)	1.00	1.33 (1.25, 1.40)	1.71 (1.62, 1.80)
Between-CCG variation (95% CI)*	0.09 (0.06, 0.12)	0.04 (0.03, 0.06)	0.01 (0.01, 0.02)	0.09 (0.07, 0.11)	0.05 (0.03, 0.06)	0.01 (0.01, 0.02)
P-value for differences in between-CCG variability**	.	< 0.01	< 0.01	.	< 0.01	< 0.01
OR for CCG at 2.5th percentile in period	0.56	0.67	0.80	0.56	0.66	0.79
OR for CCG at 97.5th percentile in period	1.77	1.49	1.24	1.79	1.52	1.27
			NSCLC			
OR for difference between time periods (95% CI)	1.00	1.18 (1.12, 1.24)	1.40 (1.33, 1.47)	1.00	1.10 (1.05, 1.16)	1.26 (1.20, 1.32)
Between-CCG variation (95% CI)*	0.04 (0.03, 0.06)	0.04 (0.03, 0.05)	0.03 (0.02, 0.05)	0.04 (0.03, 0.05)	0.03 (0.02, 0.05)	0.03 (0.02, 0.04)
P-value for differences in between-CCG variability**	.	0.84	0.61	.	0.76	0.68
OR for CCG at 2.5th percentile in period	0.68	0.69	0.70	0.69	0.70	0.70
OR for CCG at 97.5th percentile in period	1.46	1.45	1.43	1.45	1.42	1.42
			Ovarian cancer			
OR for difference between time periods (95% CI)	1.00	1.08 (0.98, 1.18)	1.12 (1.02, 1.23)	1.00	1.14 (1.02, 1.28)	1.18 (1.05, 1.31)
Between-CCG variation (95% CI)*	0.10 (0.06, 0.17)	0.04 (0.01, 0.10)	0.04 (0.02, 0.09)	0.13 (0.08, 0.23)	0.05 (0.02, 0.13)	0.05 (0.02, 0.12)
P-value for differences in between-CCG variability**	.	0.03	0.04	.	0.05	0.04
OR for CCG at 2.5th percentile in period	0.53	0.69	0.67	0.49	0.65	0.66
OR for CCG at 97.5th percentile in period	1.87	1.45	1.5	2.05	1.55	1.52

*Estimated between CCG-variance on log scale.
**Comparing between-CCG variation between 2008 – 09 and 2010 – 11; and variability between 2008 – 09 and.2012 – 13.
*** Factors adjusted for: age, sex, comorbidity status, tumour morphology, tumour topography.

which case mix is adjusted for (OR: 1.17 (95% CI: 1.05, 1.31) compared to 1.12 (95% CI: 1.02, 1.22)).

For colorectal cancer estimates were similar between the models in which case mix was and wasn't adjusted for.

3.3. Geographic inequalities

Geographic inequalities in early diagnosis decreased over time for colorectal cancer and ovarian cancer in models where case-mix was not considered (both $p < 0.05$, Table 3). Geographic inequalities for NSCLC were smaller than for the other two cancers in 2008 – 09, but there is no evidence that they decreased. Case-mix adjustment had little impact on the magnitude of inequalities, or on changes in inequalities over time.

In 2008 – 09, patients in CCGs with the lowest percentages of patients diagnosed early had 30–50% lower odds of early diagnosis compared to patients in an average CCG, even after adjustment for case mix (Table 3). By 2012 – 13 the gap had reduced to 20–30%. For colorectal cancer the reduction in CCG inequalities equate to an approximate between – CCG range in diagnosis at stages I or II of 21–46% in 2008 – 09, reducing to 38–50% in 2012 – 13 (Tables 2 and 3). For ovarian cancer they equate to a range of 16–44% in 2008 – 09 reducing to 22–40% in 2012 – 13.

4. Discussion

We report evidence for substantial increases in the percentage of patients diagnosed early in England during 2008–2013. Geographic inequalities in early diagnosis between CCGs were present in all time periods, but reduced substantially during 2008–2013 for colorectal cancer and ovarian cancer. Case-mix differences did not account for the changes we observed.

4.1. Strengths

We used multiple imputation – a gold-standard approach to minimise bias in cases where a fraction of data is irretrievably missing - in estimates of national changes in stage at diagnosis [19,20].

The patients missing stage data had poorer outcomes [16,17], were older, had more comorbidities, were more commonly diagnosed as an emergency, and less likely to receive major treatment. We estimate that excluding them leads to overstatement of early-stage diagnosis by 1–5 percentage points. For colorectal cancer and NSCLC, estimates of improvements were similar whether or not patients missing stage were included, but for ovarian cancer their exclusion led to a different conclusion of no improvement. This finding shows that patients missing recorded stage need to be considered when evaluating progress in early-stage diagnosis, to avoid bias.

We also found that emergency presentation became more common amongst patients missing stage during 2008–2013, whilst nationally it became less common. This indicates that the increase in stage recording has been skewed towards patients with better prognostic characteristics. If this trend continues, surveillance that excludes patients missing stage may become less representative of changes in the population, as patients with a recorded stage become less similar to patients without one.

We were able to exclude case-mix factors as an explanation for improvements over time. The changes in case mix also provide contextual information: the observed (unadjusted) increases in early diagnosis over time for NSCLC occurred as the case-mix skewed towards more carcinomas and increased comorbidities (factors associated with earlier diagnosis; comorbidities potentially due to incidental detection during X-ray for another condition), whilst for ovarian cancer early diagnosis increased despite a shift towards increased comorbidities and type II epithelial disease (factors associated with later diagnosis).

4.2. Limitations

We imputed missing stage information by assuming that it is missing randomly conditional on all the other information available, including on patient's subsequent survival ("missing at random", MAR). It is likely that this assumption is not entirely met, and that our approach reduced but did not eliminate bias. More work is needed to understand the mechanisms for missing data in England and evaluate how bias from it can be reduced. For example, misspecification of the imputation model could affect the magnitude of increases reported. However, given the very large effect estimates for changes in early diagnosis and geographic inequalities we found it unlikely that residual bias would change our overall conclusions.

Another restriction of this study is that data after reform of the NHS in 2013 were not available. This reform may have had a positive or negative impact on early diagnosis. Additionally, from 2010 NHS funding increases were lower in real terms than in previous years, and

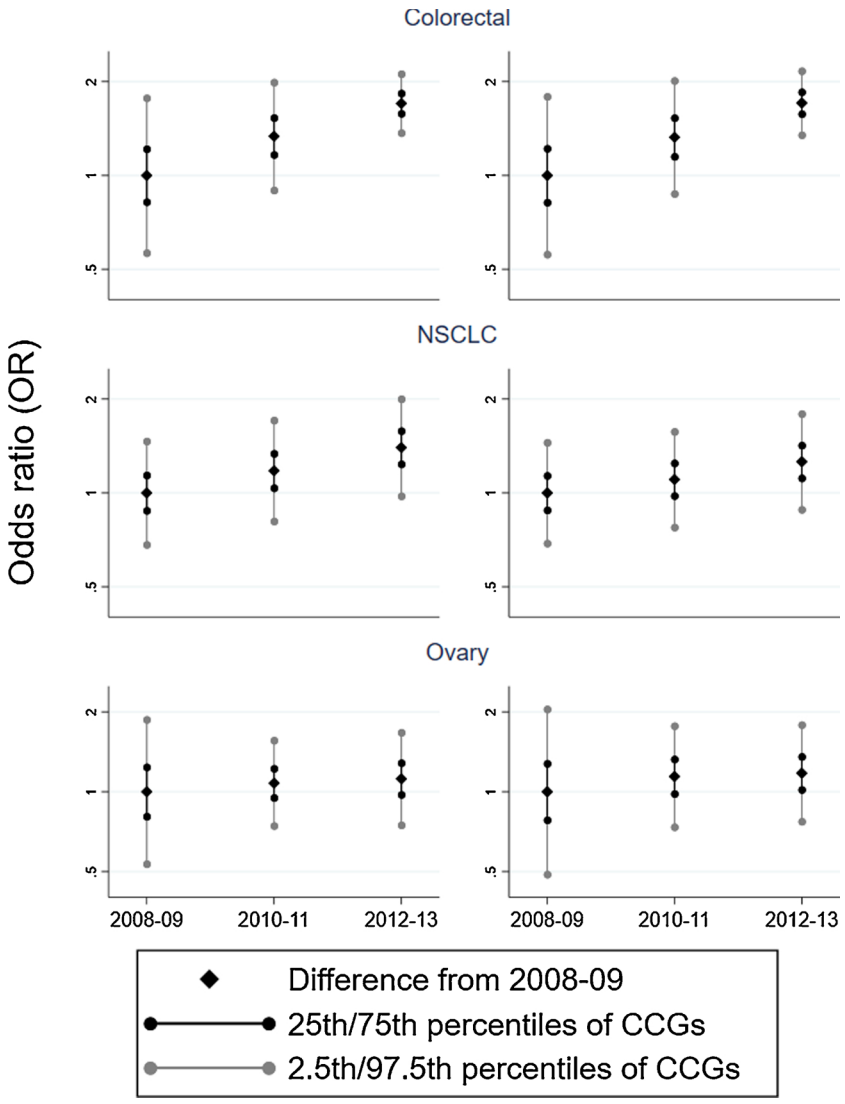


Fig. 2. Model-based estimates for change in odds of diagnosis at stages I or II, and change in between-CCG variation, in England during 2008-2013: comparison of unadjusted (left) and case-mix adjusted (right) estimates.

failed to keep pace with increases in demand, resulting in the need for efficiency savings and reductions in per-head spend on cancer by 2011 [36]. Increased waits for GP appointments and at A&E departments have also been documented [37,38], and the pressure on these gateway services may have affected early diagnosis. Our analysis cannot be used to assess the long-term impact from these changes.

4.3. Effect of early diagnosis interventions in England

The introduction of FOBT screening and increasing referrals under the TWW urgent GP referral route to diagnosis are likely to have played a role in the large increases in early diagnosis for colorectal cancer. Between 2008 – 09 and 2012 – 13 the percentage of patients diagnosed through screening rose from 6% to 10% and diagnoses through TWW rose from 26% to 30%, corresponding to 2,600 more patients per annum diagnosed through these routes (Appendix Table 5).

We estimate that by 2012 – 13 there were almost 4,000 fewer new colorectal cancer diagnoses at stage IV annually (Table 2), and corresponding increases in diagnoses at stages I, II and, to an extent, III. One concern about screening programmes is the increased risk of overdiagnosis and corresponding increase in unnecessary treatment of low-grade/benign tumours [39]. If overdiagnosis increases, there will be increases in incidence, early diagnosis, and survival without benefit to patients. Our estimates indicate that early diagnosis increases for colorectal cancer during 2008–2013 are unlikely to be due to increased overdiagnosis. This is because incidence rose only slightly, whilst the absolute number of diagnoses at stage IV dropped substantially.

For NSCLC and ovarian cancers there were also large increases in TWW diagnoses in this period. These may have resulted from the introduction of GP referral guidelines and symptom awareness campaigns.

The changes in early diagnosis during 2008–2013 occurred following sustained government investment in cancer control initiated through the National Cancer Plan in 2000 (which promised an additional £570 million for cancer by 2003 – 04) [7]. Though it is probable that the increased spending coupled with this plan (and subsequent extensions to it in 2007 and 2011 [8,11]) led to improvements in early-stage diagnosis, empirical data supporting it have thus far been sparse. Our study provides evidence for a stark improvement. It seems likely that these cancer plans have at least in part led to this, as well as contributing to reduced geographic inequalities, if national referral guidelines and targets for cancer have helped standardise patient pathways across the country.

4.4. Conclusion and recommendations

We report very large increases in the percentage of patients diagnosed at stages I or II for colorectal cancer and NSCLC during 2008–2013, and a smaller increase for ovarian cancer. The increases we report may be subject to residual bias from missing stage data, however the overall conclusion of large improvements is robust to some misestimation. Increased investment and more frequent diagnoses through screening (for colorectal cancer only) and the two-week wait route to diagnosis are likely to have contributed to the increases. Geographic inequalities reduced considerably for colorectal and ovarian cancer over the same time period.

Though useful for rapid surveillance and evaluation of success against government targets, two measures currently used by Public Health England, the "complete case" early stage percentage and missing stage percentage, give an incomplete picture of changes in early diagnosis in the population. Epidemiological analyses of stage trends are needed in addition to these in order to evaluate progress. Patient records missing stage should be included in analyses through an imputation approach as done here, or prognostic measures based on estimated stage or survival could be used [16]. This recommendation is based on analysis of patients in England, but is likely to be equally relevant to the *Detect Cancer Early* programme in Scotland [40], and other stage surveillance programmes internationally.

Our findings are based on a gold-standard approach to reduce bias when stage data are missing but auxiliary information is available. They concord with improvements in survival during this period [41]. However, further research is needed to better understand the mechanisms by which stage is missing and to optimise imputation models; to replicate our finding of a very large increase in colorectal cancer early diagnosis; and to understand the drivers of improvement.

Our analysis concludes in 2013. It is important that epidemiological analyses of trends in early stage diagnosis after this period are done in order to understand the impact of health service reform and financial austerity on cancer control.

Authorship contribution statement

PM, LW, and SW designed the study. PM did the analysis and drafted the manuscript under supervision from LW and SW. All authors contributed to the interpretation of the study results and preparing the manuscript. All authors approved the final draft for submission.

CRediT authorship contribution statement

Patrick Muller: Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing - original draft, Visualization. **Laura Woods:** Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration. **Sarah Walters:** Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

None.

Acknowledgements

PM and SW received funding from a Cancer Research UK (CRUK) Early Diagnosis Policy Research grant to the Cancer Policy Programme at the London School of Hygiene & Tropical Medicine (award number C7923/A18348), and LW is funded by a Cancer Research UK Population Research Postdoctoral Fellowship (award number C23409/A11415). The authors would like to thank attendees of the CRUK Early Diagnosis Indicators Workshop (2017) for insightful discussion regarding measurement of early diagnosis. We owe thanks to Chiara Di Girolamo, Melanie Morris, Michel Coleman, Bernard Rachet, and Sara Benitez Majano for informative discussions and support; to Matteo Quartagno for assistance and advice on the use of *jomo*; and to Helen Fowler for providing the algorithm to determine receipt of major treatment. We additionally thank Aurelien Belot and Edmund Njeru Njagi for advice on the specification of the multilevel models and analysis strategy, Adrian Turculet for performing the data cleaning and linkage, and Public Health England for providing the data used in the analysis.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.canep.2020.101743.

References

[1] Office for National Statistics, Cancer Survival by Stage at Diagnosis for England (experimental Statistics): Adults Diagnosed 2012, 2013 and 2014 and Followed up to 2015, London, UK, (2016).
[2] C. Maringe, S. Walters, B. Rachet, J. Butler, T. Fields, P.J. Finan, R. Maxwell, B. Nedrebø, L. Påhlman, A. Sjövall, A. Spigelman, G. Engholm, A. Gavin, M.L. Gjerstorff, J. Hatcher, T. Borge Johannesen, E.J. Morris, C.E. McGahan, E. Tracey, D. Turner, M.A. Richards, M.P. Coleman, ICBP Module 1 Working Group, Stage at diagnosis and colorectal cancer survival in six high-income countries: a population-based study of patients diagnosed during 2000-7, Acta Oncol. 52 (2013) 919–932.
[3] Independent Cancer Taskforce, Achieving World-Class Cancer Outcomes: A Strategy for England 2015-2020, London, UK (2015).
[4] Cancer Australia, Cancer Australia Strategic Plan 2014-2019, Surry Hills, New South Wales (2014), p. 40.
[5] Cancer Care Ontario, Ontario Cancer Plan 2008-2011, Toronto, Canada (2008).
[6] Danish National Board of Health, The National Cancer Plan Summary, Copenhagen, Denmark (2000).
[7] Department of Health, The NHS Cancer Plan, London, UK (2000).
[8] Department of Health, Improving Outcomes: A Strategy for Cancer, London, UK (2011).
[9] Scottish Government, Detect Cancer Early (programme summary). https://www2.gov.scot/Topics/Health/Services/Cancer/Detect-Cancer-Early.
[10] M.A. Richards, The size of the prize for earlier diagnosis of cancer in England, Br. J. Cancer 101 (Suppl. 2) (2009) S125–9.
[11] Department of Health, Cancer Reform Strategy, London, UK (2007).
[12] R. Roope, New NICE GP Guidelines Have Huge Ambition and Potential, (2015) https://scienceblog.cancerresearchuk.org/2015/06/23/new-nice-gp-guidelines-have-huge-ambition-and-potential/?_ga=2.101901505.1789612612.1564331422-821341162.1564331422.
[13] R.F.A. Logan, J. Patnick, C. Nickerson, L. Coleman, M.D. Rutter, C. von Wagner, Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests, Gut 61 (10) (2011) 1439–1446.
[14] Public Health England, Be Clear On Cancer - Overview, UK (2019).
[15] Public Health England, Cancer Data Dashboard, (2019) https://www.cancerdata.nhs.uk/dashboard#?tab=Overview.
[16] P. Muller, S. Walters, M.P. Coleman, L. Woods, Which indicators of early cancer diagnosis from population-based data sources are associated with short-term mortality and survival? Cancer Epidemiol. 56 (2018) 161–170.
[17] C. Di Girolamo, S. Walters, S. Benitez Majano, B. Rachet, M.P. Coleman, E.N. Njagi, M. Morris, Characteristics of patients with missing information on stage: a population-based study of patients diagnosed with colon, lung or breast cancer in England in 2013, BMC Cancer 18 (2018) 492.
[18] M.E. Barclay, G.A. Abel, L. Elliss-Brookes, D.C. Greenberg, G. Lyratzopoulos, The influence of patient case mix on public health area statistics for cancer stage at diagnosis: a cross-sectional study, Eur. J. Public Health (2019).
[19] Q. Luo, S. Egger, X.Q. Yu, D.P. Smith, D.L. O'Connell, Validity of using multiple imputation for "unknown" stage at diagnosis in population-based cancer registry data, PLoS One 12 (6) (2017) e0180033.
[20] M. Falcaro, U. Nur, B. Rachet, J.R. Carpenter, Estimating excess hazard ratios and net survival when covariate data are missing: strategies for multiple imputation, Epidemiology (Cambridge, Mass.) 26 (3) (2015) 421–428.
[21] M.E. Barclay, G. Lyratzopoulos, D.C. Greenberg, G.A. Abel, Missing data and chance variation in public reporting of cancer stage at diagnosis: cross-sectional analysis of population-based data in England, Cancer Epidemiol. 52 (Supplement C) (2018) 28–42.
[22] World Health Organisation, International Statistical Classification of Diseases and Related Health Problems - 10th Revision (ICD-10), Geneva, Switzerland (2011).
[23] Royal College of Physicians, National Lung Cancer Audit, (2015) (Accessed 03/11 2016), https://www.rcplondon.ac.uk/projects/national-lung-cancer-audit.
[24] Healthcare Quality Improvement Partnership, National Bowel Cancer Audit Report 2014, Leeds (2014).
[25] L. Elliss-Brookes, S. McPhail, A. Ives, M. Greenslade, J. Shelton, S. Hiom,

M. Richards, Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets, Br. J. Cancer 107 (8) (2012) 1220–1226.
[26] S. Benitez Majano, H. Fowler, C. Maringe, C. Di Girolamo, B. Rachet, Deriving stage at diagnosis from multiple population-based sources: colorectal and lung cancer in England, Br. J. Cancer 115 (3) (2016) 391–400.
[27] C. Maringe, H. Fowler, B. Rachet, M.A. Luque-Fernandez, Reproducibility, reliability and validity of population-based administrative health data for the assessment of cancer non-related comorbidities, PLoS One 12 (3) (2017) e0172814.
[28] J.R. Carpenter, M.G. Kenward, Multiple Imputation and Its Application, Wiley, Chichester, 2013.
[29] M. Quartagno, J. Carpenter, Jomo: a Package for Multilevel Joint Modelling Multiple Imputation, (2017).
[30] J. Gurney, D. Sarfati, J. Stanley, The impact of patient comorbidity on cancer stage at diagnosis, Br. J. Cancer 113 (9) (2015) 1375–1380.
[31] D.H. Brewster, C.S. Thomson, D.J. Hole, R.J. Black, P.L. Stroner, C.R. Gillis, Relation between socioeconomic status and tumour stage in patients with breast, colorectal, ovarian, and lung cancer: results from four national, population based studies, BMJ 322 (7290) (2001) 830–831.
[32] S. Benitez Majano, C. Di Girolamo, B. Rachet, C. Maringe, M.G. Guren, B. Glimelius, L.H. Iversen, E.A. Schnell, K. Lundqvist, J. Christensen, M. Morris, M.P. Coleman, S. Walters, Surgical treatment and survival from colorectal cancer in Denmark, England, Norway, and Sweden: a population-based study, Lancet Oncol. 20 (1) (2019) 74–87.
[33] J.W. Graham, A.E. Olchowski, T.D. Gilreath, How many imputations are really needed? Some practical clarifications of multiple imputation theory, Prev. Sci. 8 (3) (2007) 206–213.
[34] D.B. Rubin, Multiple Imputation for Nonresponse in Surveys, John Wiley and Sons, New York, 1987.
[35] S. Rabe-Hesketh, A. Skrondal, Multilevel and Longitudinal Modeling Using Stata, 3rd ed., Stata Press, 2012.
[36] Cancer Research UK, Cancer Services: Reverse, Pause or Progress? Cancer Research UK, London, 2012.
[37] More Patients Waiting Longer Than a Week for GP Appointments, The Guardian, London, 2017.
[38] I. Blunt, N. Edwards, L. Merry, What's Behind the A&E "crisis"? Nuffield Trust, London, 2015.
[39] Margaret McCartney, Patients Deserve the Truth: Heatlh Screening Can Do More Harm than Good, The Guardian, 2014.
[40] Scottish Government, Detect Cancer Early: NHS Scotland Performance against Local Delivery Plan (LDP) Standards, (2019) https://www.gov.scot/publications/nhsscotland-performance-against-ldp-standards/pages/detect-cancer-early/.
[41] S. Walters, S. Benitez Majano, P. Muller, M.P. Coleman, C. Allemani, J. Butler, M. Peake, M.G. Guren, B. Glimelius, S. Bergstrom, L. Pahlman, B. Rachet, Is England closing the international gap in cancer survival? Br. J. Cancer 113 (5) (2015) 848–860.

11.2 Supplementary appendices

Appendix Table 1 Counts of patients in each diagnosis period by CCG and stage at diagnosis

	Colorectal cancer			NSCLC			Ovarian cancer		
	2008-2009	2010-2011	2012-2013	2008-2009	2010-2011	2012-2013	2008-2009	2010-2011	2012-2013
Count of patients by CCG									
CCG Average	306	316	318	275	288	299	46	47	45
Minimum	36	43	36	62	79	57	3	8	8
Maximum	1,355	1,402	1,295	951	1,044	1,060	216	213	195
Count (%) of patients by stage at diagnosis									
I	2,480 (3.9)	5,970 (9.0)	9,899 (14.9)	5,116 (8.9)	6,879 (11.4)	9,467 (15.1)	1,041 (10.8)	1,472 (14.9)	1,979 (20.6)
II	4,859 (7.6)	10,052 (15.2)	15,074 (22.7)	2,435 (4.2)	4,219 (7.0)	5,109 (8.2)	333 (3.5)	426 (4.3)	576 (6.0)
III	6,214 (9.7)	11,247 (17.0)	15,262 (23.0)	10,893 (19.0)	11,803 (19.6)	12,995 (20.8)	1,397 (14.5)	1,875 (19.0)	3,171 (33.0)
IV	10,112 (15.8)	13,028 (19.7)	15,291 (23.0)	19,301 (33.6)	26,329 (43.8)	29,395 (47.0)	1,356 (14.1)	1,582 (16.0)	2,048 (21.3)
Missing	40,307 (63.0)	25,816 (39.0)	10,900 (16.4)	19,637 (34.2)	10,873 (18.1)	5,597 (8.9)	5,491 (57.1)	4,508 (45.7)	1,821 (19.0)

Appendix Table 2 Characteristics of the patients without a recorded stage at diagnosis

	Colorectal cancer			NSCLC			Ovarian cancer		
	2008-2009	2010-2011	2012-2013	2008-2009	2010-2011	2012-2013	2008-2009	2010-2011	2012-2013
Patients missing stage (%)	40,307 (63.0)	25,816 (39.0)	10,900 (16.4)	19,637 (34.2)	10,873 (18.1)	5,597 (8.9)	5,491 (57.1)	4,508 (45.7)	1,821 (19.0)
Average age	72	73	74	75	75	76	68	69	70
Female (%)	45.2	45.5	47.8	44.7	45.6	47.2	-	-	-
1+ comorbidities (%)	19.2	22.6	29	31.6	36.6	39.9	15	16.8	20.2
Emergency presentation (%)	23.5	24.8	29.6	46.1	48.3	50.9	36.8	37.8	40.7
Deprivation quintile 4/5 (%)	36.1	35.4	34.5	47.8	46.4	45.7	35.5	36.6	34.9

Appendix Table 3 Distribution of case-mix characteristics in each time period (with multiple imputation estimates used for topography and morphology to account for the missing data)

	Colorectal cancer				NSCLC				Ovarian cancer		
	2008-09	2010-11	2012-13		2008-09	2010-11	2012-13		2008-09	2010-11	2012-13
Age											
15-39 (%)	1.5	1.6	2.2		0.5	0.5	0.4		4.8	4.5	4.8
40-49 (%)	3.7	3.7	3.9		2.4	2.3	2.2		8.6	9.1	8.6
50-59 (%)	10.1	10.3	11.2		9.8	9.4	9.2		17.0	16.4	17.2
60-69 (%)	26.4	26.5	24.7		25.2	25.7	26.3		26.2	26.6	27.1
70-79 (%)	31.3	31.0	30.5		34.3	33.7	34.0		24.8	24.7	24.6
80-99 (%)	27.1	26.8	27.4		27.8	28.4	27.9		18.6	18.7	17.6
Sex											
Male (%)	55.7	56.0	56.3		56.5	55.8	54.7				
Female (%)	44.3	44.0	43.7		43.5	44.2	45.3				
Charlson comorbidity Index											
0 (%)	81.9	80.1	77.7		71.7	68.6	65.9		87.0	85.5	84.4
>0 (%)	18.1	19.9	22.3		28.3	31.4	34.1		13.0	14.5	15.6
Topography											
Colon (%)	64.3	64.6	65.2	Main bronchus	9.0	7.6	6.1	Ovary	97.7	97.2	95.8
Rectum (%)	35.7	35.4	34.8	Lobe	91.0	92.4	93.9	Fallopian tube	2.3	2.8	4.2
Morphology											
Carcinoma (%)	90.2	89.9	88.2	Carcinoma	49.5	57.1	59.3	Type I epithelial	19.4	18.9	18.9
Non-carcinoma (%)	9.8	10.1	11.8	Non-carcinoma	50.5	42.9	40.7	Type II epithelial	77.2	78.0	78.2
								Non-epithelial	3.4	3.1	2.9

Appendix Table 4 Multi-level logistic regression results: Odds ratios (and 95% CIs) for association between patient factors and probability of diagnosis at stages I or II

	Colorectal cancer		NSCLC		Ovarian cancer	
Diagnosis period						
2008-2009	1.00		1.00		1.00	
2010-2011	1.33 (1.25, 1.40)		1.10 (1.05, 1.16)		1.14 (1.02, 1.28)	
2012-2012	1.71 (1.62, 1.80)		1.26 (1.20, 1.32)		1.18 (1.05, 1.31)	
Age						
15-39	1.00		1.00		1.00	
40-49	0.90 (0.82, 0.98)		0.36 (0.30, 0.43)		0.74 (0.58, 0.93)	
50-59	1.02 (0.95, 1.10)		0.35 (0.30, 0.41)		0.61 (0.50, 0.75)	
60-69	1.29 (1.20, 1.39)		0.42 (0.35, 0.49)		0.45 (0.37, 0.55)	
70-79	1.30 (1.21, 1.39)		0.43 (0.36, 0.50)		0.36 (0.29, 0.44)	
80-99	1.05 (0.97, 1.13)		0.40 (0.34, 0.47)		0.27 (0.22, 0.34)	
Sex						
Male	1.00		1.00		NA	
Female	1.00 (0.98, 1.02)		1.19 (1.19, 1.19)			
Charlson comorbidity Index						
0	1.00		1.00		1.00	
1	1.06 (1.02, 1.09)		1.55 (1.50, 1.60)		0.93 (0.81, 1.07)	
2	0.97 (0.93, 1.01)		1.52 (1.45, 1.59)		1.04 (0.87, 1.23)	
3+	1.06 (1.02, 1.11)		1.55 (1.48, 1.62)		0.69 (0.52, 0.91)	
Topography*						
Colon	1.00	Lobe	1.00	Ovary	1.00	
Rectum	1.12 (1.10, 1.14)	Main Bronchus	0.25 (0.23, 0.28)	Fallopian tube	3.37 (2.84, 4.01)	
Morphology						
Carcinoma	1.00	Carcinoma	1.00	Type I epithelial	1.00	
Non-carcinoma	2.10 (2.03, 2.16)	Non-carcinoma	0.60 (0.58, 0.62)	Type II epithelial	0.07 (0.07, 0.08)	
				Non-epithelial	0.58 (0.46, 0.74)	

* Cancer sites are defined by the following ICD codes: Colon=C18; Rectum=C19,C20,C21.8; Main Bronchus=C34.0; Lobe=C34.1,C34.2,C34.3; Ovary=C56; Fallopian tube=C57

Appendix Table 5 Counts of patient's route to diagnosis, cancer, period of diagnosis, 2008-2013

	Colorectal cancer			NSCLC			Ovarian cancer		
Route to diagnosis (%)	2008-09	2010-11	2012-13	2008-09	2010-11	2012-13	2008-09	2010-11	2012-13
Emergency presentation	14,743 (23.1)	14,563 (22.0)	15,325 (23.1)	21,230 (37.0)	21,284 (35.4)	21,617 (34.6)	3,133 (32.6)	3,030 (30.7)	2,845 (29.7)
GP referral	15,343 (24.0)	15,120 (22.9)	15,297 (23.0)	12,135 (21.2)	12,326 (20.5)	13,446 (21.5)	2,173 (22.6)	2,077 (21.1)	2,069 (21.6)
Inpatient Elective	2,841 (4.4)	2,322 (3.5)	2,174 (3.3)	937 (1.6)	894 (1.5)	893 (1.4)	161 (1.7)	114 (1.2)	91 (1.0)
Other outpatient	5,135 (8.0)	4,588 (6.9)	4,304 (6.5)	6,050 (10.5)	6,258 (10.4)	7,202 (11.5)	1,156 (12.0)	1,104 (11.2)	912 (9.5)
Screening	3,738 (5.8)	6,415 (9.7)	6,404 (9.6)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
TWW	16,502 (25.8)	17,981 (27.2)	19,766 (29.8)	13,688 (23.9)	15,340 (25.5)	17,619 (28.2)	2,316 (24.1)	2,874 (29.1)	3,410 (35.5)
Unknown or missing	5,670 (8.9)	5,124 (7.8)	3,156 (4.8)	3,342 (5.8)	4,001 (6.7)	1,786 (2.9)	679 (7.1)	664 (6.7)	268 (2.8)

Appendix Table 6: List of OPCS codes used to determine receipt of major surgical treatment. This list was produced by staff in the Cancer Survival Group at LSHTM and kindly provided to the authors by Helen Fowler.

Code	Description
	Colorectal cancer
H011	Emergency excision of abnormal appendix and drainage HFQ
H012	Emergency excision of abnormal appendix NEC
H013	Emergency excision of normal appendix
H018	Other specified emergency excision of appendix
H019	Emergency appendicectomy NEC, unspecified
H021	Interval appendicectomy
H022	Planned delayed appendicectomy NEC
H023	Prophylactic appendicectomy NEC
H024	Incidental appendicectomy
H028	Other specified other excision of appendix
H029	Appendicectomy NEC, unspecified;
H041	Proctocolectomy NEC, Panproctocolectomy and Ileostomy
H042	Panproctocolectomy and anastomosis of ileum to anus and creation of pouch HFQ
H043	Panproctocolectomy and anastomosis of ileum to anus NEC
H048	Other specified total excision of colon and rectum
H049	Panproctocolectomy NEC, Total excision of colon and rectum, unspecified-
H051	Total colectomy and anastomosis of ileum to rectum
H052	Total colectomy and ileostomy and creation of rectal fistula HFQ
H053	Total colectomy and ileostomy NEC
H058	Total excision of colon, other specified
H059	Total excision of colon, Unspecified
H061	Extended right hemicolectomy and end to end anastomosis
H062	Extended right hemicolectomy and anastomosis of ileum to colon
H063	Extended right hemicolectomy and anastomosis NEC
H064	Extended right hemicolectomy and ileostomy HFQ
H068	Other specified extended excision of right hemicolon
H069	Extended excision of Right hemicolon, unspecified, excision of Right colon and surrounding tissue
H071	Right hemicolectomy and end to end anastomosis of ileum to colon, Ileocaecal resection
H072	Right hemicolectomy and side to side anastomosis of ileum to transverse colon,
H073	Right hemicolectomy and anastomosis NEC
H074	Right hemicolectomy and ileostomy HFQ
H078	Other specified other excision of right hemicolon
H079	Other excision of right hemicolon, unspecified; Right hemicolectomy NEC
H081	Transverse colectomy and end to end anastomosis
H082	Transverse colectomy and anastomosis of ileum to colon
H083	Transverse colectomy and anastomosis NEC
H084	Transverse colectomy and ileostomy HFQ
H085	Transverse colectomy and exteriorisation of bowel NEC (CODE COLOSTOMY SPERATELY)
H088	Other specified excision of transverse colon
H089	Excision of transverse colon, unspecified
H091	Left hemicolectomy and end to end anastomosis of colon to rectum
H092	Left hemicolectomy and end to end anastomosis of colon to colon
H093	Left hemicolectomy and anastomosis NEC
H094	Left hemicolectomy and ileostomy HFQ
H095	Left hemicolectomy and exteriorisation of bowel NEC (CODE COLOSTOMY SEPERATELY)
H098	Excision of left hemicolon, Other specified
H099	Left hemicolectmy NEC, Excision of left hemicolon, Unspecified
H101	Sigmoid colectomy and end to end anastomosis of ileum to rectum
H102	Sigmoid colectomy and anastomosis of colon to rectum
H103	Sigmoid colectomy and anastomosis NEC
H104	Sigmoid colectomy and ileostomy HFQ
H105	Sigmoid colectomy and exteriorisation of bowel NEC
H108	Other specified excision of sigmoid colon
H109	Unspecified excision of sigmoid colon
H111	Colectomy and end to end anastomosis of colon to colon NEC
H112	Colectomy and side to side anastomosis of ileum to colon NEC
H113	Colectomy and anastomosis NEC
H114	Colectomy and ileostomy NEC

Code	Description
	Colorectal cancer (continued)
H115	Colectomy and exteriorisation of bowel (CODE COLOSTOMY SEPERATELY)
H118	Other excision of colon, other specified
H119	Hemicolectomy NEC; Colectomy NEC, Other excision of colon, unspecified;
H121	Excision of diverticulum of colon
H122	Polypectomy NEC, Excision of lesion NEC
H123	Destruction of lesion of colon NEC
H128	Other specified extirpation of lesion of colon
H129	Unspecified extirpation of lesion of colon
H291	Subtotal excision of colon and rectum and creation of colonic pouch and anastomosis of colon to anus
H292	Subtotal excision of colon and rectum and creation of colonic pouch NEC
H293	Subtotal excision of colon and creation of colonic pouch and anastomosis of colon to rectum
H294	Subtotal excision of colon and creation of colonic pouch NEC
H298	Subtotal excision of colon, Other specified
H299	Subtotal excision of colon, Unspecified
H331	Abdominoperineal excision of rectum and end colostomy; APR; SCAPER
H332	Proctectomy and anastomosis of colon to anus
H333	Anterior resection of rectum and anastomosis of colon to rectum using staples
H334	Anterior resection of rectum and anastomosis NEC
H335	Hartmann procedure, Rectosigmoidectomy and closure of rectal stump and exteriorisation of bowel (CODE COLOSTOMY SEPERATELY)
H336	Anterior resection of rectum and exteriorisation, (CODE COLOSTOMY SEPARATELY)
H337	Perineal resection of rectum HFQ
H338	Anterior Resection of Rectum NEC, Rectosigmoidectomy and anastomosis of colon to rectum Excision of rectum, other specified
H339	Rectosigmoidectomy NEC, Excision of rectum, unspecified;
H341	Open excision of lesion of rectum: Open removal of polyp; Yorke Mason
H342	Open cauterisation of lesion of rectum, Diathermy
H343	Open cryotherapy to lesion of rectum
H344	Open laser destruction of lesion of rectum
H345	Open destruction of lesion of rectum NEC
H348	Open removal of lesion of rectum, other specified
H349	Open removal of lesion of rectum, unspecified
H401	Trans-sphincteric excision of mucosa of rectum
H402	Trans-sphincteric excision of lesion of rectum
H403	Trans-sphincteric destruction of lesion of rectum
H404	Trans-sphincteric anastomosis of colon to anus
H408	Other specified operations on rectum through anal sphincter
H409	Unspecified operations on rectum through anal sphincter
X141	Total exenteration of pelvis
X142	Anterior exenteration of pelvis
X143	Posterior exenteration of pelvis
X148	Other specified clearance of pelvis
X149	Clearance of pelvis, unspecified
	NSCLC
E391	Open excision of lesion of trachea
E398	Other specified partial excision of trachea
E399	Unspecified partial excision of trachea
E441	Excision of carina
E461	Sleeve resection of bronchus and anastomosis HFQ
E541	Total pneumonectomy, total removal of lung, Pneumonectomy NEC
E542	Bilobectomy of lung
E543	Lobectomy of lung
E544	Excision of segment of lung
E545	Partial lobectomy of lung NEC
E548	Excision of lung, other specified
E549	Excision of lung, Unspecified
E552	Open excision of lesion of lung
E559	Open removal of lesion of lung, unspecified
T013	Excision of lesion of chest wall
T023	Insertion of prothesis into chest wall NEC

Code	Description
	Ovarian cancer
Q071	Radical Hysterectomy (removes uterus + cervix + vagina). Wertheims hysterectomy
Q072	Abdominal Hysterectomy and excision of periuterine tissue NEC.Radical Hysterectomy
Q074	TAH, Panhysterectomy, hysterectomy NEC (removes uterus + cervix). Total abdominal hysterectomy NEC
Q221	Bilateral salpingoophorectomy
Q223	Bilateral oophorectomy, excision of gonads
Q231	Unilateral salpingoophorectomy NEC
Q232	Salpingoophorectomy of remaining solitary fallopian tube and ovary
Q235	Unilateral oophorectomy NEC
Q241	Salpingoophorectomy NEC
Q243	Oophorectomy NEC
Q431	Excision of wedge of ovary
Q432	Excision of lesion of ovary - cystectomy
Q433	Marsupialisation of lesion of ovary
Q439	Unspecified partial excision of ovary
Q441	Open cauterisation of lesion of ovary
Q449	Unspecified open destruction of lesion of ovary
Q478	Other specified other open operations on ovary
Q479	Unspecified other open operations on ovary
Q498	Other specified therapeutic endoscopic operations on ovary
Q499	Unspecified therapeutic endoscopic operations on ovary
Q518	Other operations on ovary, Other specified
Q519	Other operations on ovary, Unspecified
T361	Omentectomy – Complete
T865	Para-aortic lymph node sampling
T868	Pelvic lymph node sampling
T875	Para-aortic lymphadenectomy
T878 + Z941	Bilateral pelvic lymphadenectomy
T878 + Z942	Right pelvic lymphadenectomy
T878 + Z943	Left pelvic lymphadenectomy
X141	Total exenteration of pelvis
X142	Anterior exenteration of pelvis
X143	Posterior exenteration of pelvis
X148	Clearance of Pelvis OS
X149	Clearance of Pelvis unspecified

11.3 Further analysis and interpretation

The analyses presented in this paper provide evidence for a very large increase in early diagnosis for colorectal cancer and NSCLC during the period 2008-2013, and evidence that geographic inequalities reduced during that period. These results aid evaluation of the impact of health service changes during 2008-2013. It is likely introduction of colorectal screening, initiatives to expedite referrals and raise awareness of symptoms, and sustained investment in cancer to 2010, played a significant role in improvements. There was no evidence that reductions in per capita health spend, or changes in the structure of health services, up until 2013 had any negative effect on early diagnosis sufficient to negate the effect of concurrent improvements.

The analysis has several potential limitations. A binary stage I/II vs III/IV indicator was used for the substantive analysis, as a simpler alternative to a multinomial logistic regression. However, it was highlighted in Chapter 7 that each incremental difference in stage from I to IV is important for survival. Multiple imputation estimates of the granular stage distribution (I, II, III, and IV) in each year were therefore also produced, which allowed appraisal of stage shifts from II to I and IV to III during 2008-2013.

The key assumption of the analysis is that, on average, stage at diagnosis is the same for patients with and without recorded stage who are the same with respect to the variables included in the imputation model. There is potential for residual bias due to this assumption, of an assumed small but unknown magnitude. Other confirmatory analyses, using different techniques for analysis and handling missing data, would be valuable to validate the findings of large increases in early diagnosis for colorectal cancer and NSCLC.

As discussed in section 10.1.3, an *a priori* conceptual justification of the imputation approach taken can be made, with the caveat that some bias may be persist due to missing grade. Though not included in the substantive analysis models, for colorectal cancer only there is sufficient data on grade for patients missing stage (66.7% complete) that its inclusion could impact results in the imputation model. One posthoc sensitivity analysis was done after the substantive analysis, including grade in the imputation model in addition to morphologic cell type category. That analysis returned an estimate of early-stage diagnosis increasing from 31.8% in 2008-09 to 43.9% in 2012-13 (compared to 31.9% to 43.8% in the substantive analysis), confirming that no bias was introduced from failing to include grade in that imputation model.

To the author's knowledge this is the first study to use multiple imputation to evaluate national trends in stage at diagnosis, in any country, and comparable data from other settings is not immediately available. England was one of the first countries to introduce population-wide colorectal cancer screening, and only initial results from other comparable screening programmes are available. Data from the Danish programme initiated in 2014 do indicate a step-change in early diagnosis as a result [177], and evidence from the Australian programme which was gradually implemented from 2006 is similarly positive [178]. However, the evaluations of these programmes did not assess stage trends in the whole population. The rapid expansion of the English screening programme from 2006 is likely to

be one of the drivers of the improvement in early-stage diagnosis. However, the total percentage of colorectal patients diagnosed following screening rose to only 10% of all patients in 2010 [179], representing a moderate shift in how the disease is diagnosed. In addition to the rollout of screening, there may have been increased use of technologies such as colonoscopy or flexible sigmoidoscopy outside the programme.

Survival data can also be used to assess the plausibility of the reported increase in early diagnosis. There was a step-change increase in survival during 2008-2013 in England for each of the cancers analysed, consistent with a large increase in early diagnosis [75]. For lung cancer, Walters *et al* (2015) report an acceleration in survival which is even larger than the increase in early-stage diagnosis estimated. In that analysis the 5-year survival estimates for 2013 were calculated using a hybrid approach, using follow up in 2013 of patients diagnosed in earlier years [180]. In the context of increasing survival over time, the hybrid approach may underestimate the actual improvement to 2013, as the survival after one year for patients diagnosed in 2013 is estimated based on the survival in equivalent periods of follow up of patients diagnosed in earlier years, which may be lower.

Some basic scenario modelling for the impact of the stage trends on survival for colorectal cancer can be conducted as a sense-check of the early diagnosis increase reported, as follows:

- Taking estimates of stage-specific colorectal cancer 1-year net survival in England in 2010-12 for patients with stage recorded only: 98.7% at stage I; 94.8 at stage II, 88.5% at stage III, and 51.7% at stage IV [47];
- Making the strong assumption that stage-specific net survival remained static at these levels during 2008-2013;
- Estimating the changes in survival that would result between 2008-09 and in 2012-13 with these stage-specific survival estimates and the stage distributions estimated in the substantive analysis.

The above calculation results in an increase in survival of 6.1%. This is slightly higher than, but comparable to the ~4.5% net survival increase between 2008 and 2013 reported by Walters *et al*. It illustrates that the increase in early diagnosis reported in the substantive analysis would be expected to be accompanied by a smaller change in survival. Additionally, for colorectal cancer, any estimate of the impact of the stage shift is very sensitive to the (true) differences in 1-year net survival between stages I-III and stage IV, which are not known because of bias in stage-specific survival estimates from the exclusion of patients missing a recorded stage in the above survival estimates.

In summary, the substantive thesis analysis results provide evidence for a large improvement in early diagnosis. The results are concordant with the limited other data available. Further analyses to examine potential bias from the multiple imputation approach taken to handle missing data are needed, however. Further examination is particularly important for colorectal cancer and NSCLC, given the magnitude of the early diagnosis increase reported.

12. Multiple imputation for missing stage information in estimates of colorectal cancer early diagnosis trends in England: a population-based study

In this Chapter different sensitivity analyses of the multiple imputation approach used to assess colorectal cancer early diagnosis trends are presented. A simulation is used to test the imputation model accuracy in one plausible scenario. The impact of including different variables and parameters in the imputation model on the estimates of early diagnosis trends is also measured.

The extent of stage misclassification needed to change the substantive analysis conclusion about trends for colorectal cancer is then estimated using a pattern-mixture sensitivity analysis. To conduct this analysis, for each of the datasets where stage has been imputed if missing, a fraction of the imputed stage values are manually changed before the analysis model is run. The direction of the change, and exact percentage of values changed, were chosen by me to allow evaluation of how much error there would need to be in the imputation for the analysis conclusions to change substantially. A potential mechanism for error in the estimates presented in Chapter 11 is that my model consistently incorrectly imputed missing stage to later stages it actually was, and as a result over-estimated the early diagnosis improvement between 2008-09 (when missing stage data was common) and 2012-13 (when it was rarer). To explore this, in the analysis that follows I manually change up to 50% of the missing stage values which the imputation model imputed to III/IV to be I/II instead through random selection of imputed values. I find the exact percentage needed to be changed to I/II for the odds ratio for the early diagnosis changes over time in the analysis model to go from above 1 (indicating an increase in early diagnosis) to below 1 (indicating a decrease in early diagnosis). This exact percentage represents the proportion of all missing values that I imputed to late-stage disease which would have to have been misclassified (i.e. actually be early-stage disease) in order for the analysis conclusions of an improvement in early diagnosis over time to change.

Following the pattern-mixture analysis, the robustness of the imputation results are then considered with reference to patient's survival, and the literature on multiple imputation and missing stage data.

This project addresses the fourth thesis aim: to assess the potential bias from missing data in the thesis substantive analysis. In considering the optimal approach for handling missing data, the results also contribute to the final aim: to provide recommendations for future monitoring of early diagnosis. The analysis has been submitted to the *Journal of Epidemiology and Community Health* as a research report. The submission is presented here in full together with appendices and further corrections applied after the initial submission, followed by further discussion of the implications of the results for early diagnosis monitoring.

12.1 Submitted Manuscript

London School of Hygiene & Tropical Medicine
Keppel Street, London WC1E 7HT

T: +44 (0)20 7299 4646
F: +44 (0)20 7299 4656
www.lshtm.ac.uk

RESEARCH PAPER COVER SHEET

Please note that a cover sheet must be completed for each research paper included within a thesis.

SECTION A – Student Details

Student ID Number	270040	**Title**	Mr
First Name(s)	Patrick		
Surname/Family Name	Muller		
Thesis Title	Statistical approaches for monitoring early cancer diagnosis in England		
Primary Supervisor	Dr Laura Woods		

If the Research Paper has previously been published please complete Section B, if not please move to Section C.

SECTION B – Paper already published

Where was the work published?			
When was the work published?			
If the work was published prior to registration for your research degree, give a brief rationale for its inclusion			
Have you retained the copyright for the work?*	Choose an item.	Was the work subject to academic peer review?	Choose an item.

*If yes, please attach evidence of retention. If no, or if the work is being included in its published format, please attach evidence of permission from the copyright holder (publisher or other author) to include this work.

SECTION C – Prepared for publication, but not yet published

Where is the work intended to be published?	Journal of Epidemiology and Community Health
Please list the paper's authors in the intended authorship order:	Patrick Muller, Laura Woods
Stage of publication	**Submitted**

SECTION D – Multi-authored work

For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)	I was the lead author of this study. I planned the study, conducted the analysis, and prepared the draft of the paper. Co-authors provided feedback and input on the study design and the manuscript draft.

SECTION E

Student Signature	P Muller
Date	10/05/2021

Supervisor Signature	L Woods
Date	10/05/2021

Multiple imputation to minimise bias from missing stage information in estimates of early cancer diagnosis in England: a population-based study

ABSTRACT

Introduction

Monitoring early diagnosis is a priority of cancer policy in England. Information on stage has not always been available for a large proportion of patients, however, which may bias evaluations, particularly temporal comparisons. We previously estimated that early-stage diagnosis of colorectal cancer rose from 32% to 44% during 2008-2013, using a multiple imputation model. Here we examine the underlying assumptions of that model to assess its robustness.

Methods

Individually-linked cancer registration, Hospital Episode Statistics (HES), and audit data were examined. Six imputation models including different interaction terms, post-diagnosis treatment and survival information were assessed, and comparisons drawn with the *a priori* optimal model. Models were further tested by setting stage values to missing for some patients under one plausible mechanism, then comparing actual and imputed stage distributions for these patients. Finally, a pattern-mixture sensitivity analysis was conducted.

Results

Data from 196,511 colorectal patients were analysed, with 39.2% missing stage. Inclusion of survival time increased the accuracy of imputation: the odds ratio for change in early-stage diagnosis during 2008-2013 was 1.7 (95% CI: 1.6, 1.7) with survival to 1 year included, compared to 1.9 (95% CI 1.9-2.0) with no survival information. Imputation estimates of stage were accurate in one plausible simulation. Pattern-mixture analyses indicate our previous analysis conclusions would only change materially if stage were misclassified for 20% of the patients who had it categorised as late.

Conclusions

Multiple imputation models can substantially reduce bias from missing stage, but data on patient's one-year survival should be included for highest accuracy.

What is already known on this subject?

Evaluation of progress in early diagnosis is a priority of cancer policy, but assessments of the proportion of patients diagnosed at an early stage can be biased by missing stage data. Multiple imputation can reduce bias from missing stage data by using auxiliary information on patients, such as their survival time, to impute stage values.

What this study adds?

We show that including data on survival time improves the performance of multiple imputation models for stage, and also that imputation models return accurate estimates in one scenario where missing stage is associated with patient survival and period of diagnosis. Pattern-mixture analyses indicate our previous finding of a large increase in early diagnosis for colorectal cancer during 2008-2013 is robust. Our results, together with conceptual considerations about the mechanisms for missing data and evidence from other studies, confirm that multiple imputation is the ideal analytical approach to correct for missing stage information in assessments of population-level changes in early diagnosis.

INTRODUCTION

Cancer registration, the collation of data on all new cancer diagnoses, is the bedrock of cancer intelligence. In England, there have been improvements in the range and completeness of information recorded over recent decades [1]. Recording of the disease stage at diagnosis, the extent to which the primary tumour has spread beyond its original location of origin, was historically poor but improved greatly after 2005. More than one third (37%) of cancer registrations for patients diagnosed in 2008-09 had stage recorded, rising to 84% in 2012-13 and higher percentages in more recent years [1]. This improvement makes accurate evaluations of national differentials and temporal trends in early-stage diagnosis increasingly viable. However, patients without stage recorded in their registration (either because it was not ascertained at the hospital or not recorded) still need to be accounted for in the analysis. These patients have poorer outcomes than patients with stage recorded [2], partly because acutely unwell patients are less likely to have a full staging investigation completed. Simply excluding them can introduce substantial bias, and overstate success in early diagnosis. Current practice for routine surveillance is to exclude these patients. However, with robust methods to account for any bias from the missing stage data, the full impact of key events or interventions on the stage distribution of newly diagnosed cases can now be assessed. For example, the impact on early diagnosis from the nationwide introduction of colorectal cancer screening from 2007 [3], the implementation of the Health and Social Care Act 2012 in England from 2013 [4], or the COVID-19 pandemic from 2020 [5].

One statistical approach to minimise bias arising from missing values is multiple imputation (MI) [6]. MI uses auxiliary information on the patients whose stage information is missing to impute a likely distribution for stage for each patient missing it using a statistical model, then sampling repeatedly from this distribution to create *m* datasets where data is complete for every patient. Estimates of the parameter of interest are calculated using each of the *m* complete datasets: an average parameter estimate is generated along with confidence intervals which account for the increased uncertainty

arising from the imputation [7]. Imputation has been shown to be less biased than either a complete-case analyses or a 'missing value' analysis, because it makes more plausible assumptions about the missing data [6]. However, the approach assumes stage is missing randomly conditional only on the auxiliary variables used in the imputation model ("missing at random" or MAR), and that the relationships between variables specified in the imputation model match the actual (unknown) relationships in the data. If either of these assumptions is broken some residual bias in the estimates may be present.

We previously estimated stage trends during 2008-2013 in England for colorectal cancer using MI, reporting evidence for a very large increase in early-stage diagnosis [8]. Here we revisit that analysis, and empirically examine the two underlying assumptions made: that the imputation model is correctly specified, and that the data are MAR. For the first assumption we compare estimates between models of different complexity and which include different auxiliary variables. We then test whether the models correctly estimate stage trends under a plausible MAR scenario. For the assumption that the stage data are MAR itself, we perform a pattern-mixture sensitivity analysis [9]: estimating the proportion of missing values which would have to have been imputed incorrectly for conclusions from our previous analysis to change. We consider the plausibility of this, with reference to one year survival after diagnosis and consideration of likely mechanisms for missing data. The implications of our findings for the surveillance are then considered with reference to the early diagnosis targets set in NHS long Term Plan (2019) [10].

METHODS

Data sources and items have been described previously [8]. In brief, data from the National Cancer Registry on the 196,511 patients diagnosed with colorectal cancer in England during the period 2008-2013 were individually linked to data from national clinical audits and the Hospital Episodes Statistics (HES) database. This allowed us to obtain information on TNM stage [11], tumour topography and morphology, routes to diagnosis [12], pre-existing comorbidities and receipt of major treatment within 3 months following diagnosis [8, 13] for each person. Data on patient's date of death from the NHS Central Register were also linked. Data were complete for all patients up to 31 December 2014, hence length of vital status follow up ranged from a maximum of 7 years (patients diagnosed on 1 January 2008) to a minimum of 1 year (patients diagnosed on 31 December 2013).

We examined two separate issues for the imputation model specification: i) whether data on the survival time from diagnosis and ii) whether interactions between time period and other key predictive variables improved the accuracy of the imputation. The first issue concerns how accurate a very timely multiple imputation analyses of stage trends could theoretically be, whilst the latter concerns whether bias is introduced from assuming a homogenous relationship between different predictor variables and stage (e.g. that survival time has a consistent relationship with stage) at different times within the study period. The models were run treating stage as an unordered categorical variable with four possible values (I – IV), using the R package *jomo* which accounted for the clustering (similarity

in outcome) of patients diagnosed and treated within the same Clinical Commissioning Group (CCG) area [14]. This facilitated use of analysis models with a multilevel structure, which account for the expected clustering and hence return more accurate parameter estimates and confidence intervals. Multilevel logistic regression analysis models were then fitted using the Stata package meqrlogit [15], with stage dichotomised into early (TNM I or II) or late (III or IV).

Phase 1

First, imputation models using different combinations of auxiliary variables were compared (Table 1). The models were used to estimate changes in early-stage diagnosis during 2008-2013. All the models included information captured up to the point of diagnosis: age at diagnosis, sex, CCG of residence, quarter year of diagnosis, deprivation, comorbidities, route to diagnosis (of eight possible, including emergency presentation and GP referral), and topographical sub-site (colon or rectum). Models A and B included only this pre-diagnostic information, with model B including interactions between time period of diagnosis and each of age, sex, and topography. Models C and D included post-diagnostic information on survival truncated at one year after diagnosis (parameterised as cumulative hazard and a vital status indicator) and receipt of major surgical treatment within 3 months following diagnosis. Model D included the same interactions as Model B, and additional interaction terms between the cumulative hazard and time period. Models E and F were specified identically as models C and D, respectively, but included post-diagnosis information on survival up to end of follow for all patients (to 31 December 2014). As the most complex model, F was considered *a priori* the most accurate and other models were evaluated based on the extent to which results from them deviated from it. The issue of accuracy of very timely multiple imputation analyses was evaluated by comparison between models A, C, and E; and evaluation of bias from assuming a homogenous relationship between different predictor variables and stage in different calendar years was evaluated by comparisons between A and B, between C and D, and between E and F.

Phase 2

In phase 2, only patients with disease stage recorded were selected for analysis. A sample of 40% of these patients then had stage recoded to missing, with patients diagnosed in early years and (separately) who died shortly after diagnosis more likely to have stage missing (code in Appendix 1). This is a plausible MAR missingness mechanism, as stage was less well recorded in earlier time periods, and is more likely missing for acutely unwell and elderly patients [2]. Imputation model E, which used all available information on the patients, was then used to estimate stage for the 40% with stage removed on the basis of information on the remaining 60% of patients. The percentages of patients with each stage, and change in the stage I/II percentage, were then compared to the known values.

Table 1. Auxiliary variables included in the imputation models

Auxiliary variables included	**Model**					
	(A)	(B)	(C)	(D)	(E)	(F)
Age	X	X	X	X	X	X
Sex	X	X	X	X	X	X
CCG	X	X	X	X	X	X
Quarter year[1]	X	X	X	X	X	X
Deprivation	X	X	X	X	X	X
Charlson score	X	X	X	X	X	X
Route to diagnosis	X	X	X	X	X	X
Tumour topography	X	X	X	X	X	X
*Time period[1] * Age interaction*		X		X		X
*Time period[1] * Sex interaction*		X		X		X
*Time period[1] * Topography interaction*		X		X		X
*Time period[1] * emergency presentation interaction*		X		X		X
Major cancer treatment within 3 months of diagnosis			X	X	X	X
One year only: Cumulative hazard from diagnosis to end of follow up			X	X		
One year only: Dead or alive at end of follow up indicator			X	X		
One year only: *Time period[1] * cumulative hazard interaction*				X		
Up to six years: Cumulative hazard from diagnosis to end of follow up					X	X
Up to six years: Dead or alive at end of follow up indicator					X	X
Up to six years: *Time period[1] * cumulative hazard interaction*						X

[1] Quarter year of diagnosis indicates which of the 3-month periods during 2008-2013 the patient was diagnosed in (of 24 total). For parsimony with variables, interaction terms between time and other variables were fitted using two-year time periods of diagnosis (2008-09, 2010-11, or 2012-13).

Phase 3

In the final phase, a pattern-mixture sensitivity analysis was employed to assess the robustness of our previous conclusions, specifically evaluating to what extent the MAR assumption would need to be breached to explain the large increases in early-stage diagnosis we previously observed. One scenario for error is that previously the missing stage was imputed to later stages than it should have been, leading to overestimation of early diagnosis increases during a period when recording of stage also improved. To assess how common this misclassification would have to be for analysis conclusions to change, a varying percentage of the patients missing stage (α%), originally estimated to have stage III/IV, were randomly selected and recoded to stage I/II. Multiple imputation estimates were then re-run. Through repeated testing, the value of α needed to reduce the estimate of the increase in early diagnosis by half as well as to 0 percentage points was determined. The one-year (overall, crude) survival of the patients by recorded and imputed stage was also estimated and used to help consider the plausibility of such a scenario.

RESULTS

Phase 1

Of the 196,511 patients diagnosed with colorectal cancer in England during 2008-2013, 77,023 (39.2%) did not have a record of stage at diagnosis. A higher percentage of patients had missing stage data in 2008-09 (63.0%) than in 2013-13 (16.4%). Stage was most frequently missing for patients who died shortly after diagnosis, for older patients, for patients with pre-existing comorbidities, for patients missing information on tumour morphology, and for patients diagnosed following emergency presentation.

The inclusion of information on treatment and survival time after diagnosis in these data had a noteworthy impact on the accuracy of the estimated stage distribution throughout 2008-2013 (Table 2). For example, early-stage diagnosis during 2008-09 was estimated at 29.4% where survival data was not utilised (model A) compared to 31.9% where it was (Model E). The odds ratio for the increase in early diagnosis during 2008-2013 was also erroneously slightly larger when survival time was not used: 1.9 (95% CI: 1.9-2.0) in Model A compared to 1.7 (95% CI: 1.6-1.7) in Model E.

When information on survival was restricted to 12 months (models (C) and (D)), estimates for the earliest period 2008-09 were closer to those with no survival time information than to the optimal model (for example, C: 30.1%; 95% CI: 29.4-30.9 compared to A: 29.4%; 95%CI: 28.7-30.1). Results for the periods 2010-11 and 2012-13 from model (E) were, by contrast, very close to the optimal model: 38.1% (95%CI: 37.6-38.6) compared to 38.5% (95%CI: 38.0-39.0) and 43.7% (95%CI: 43.2-44.1) compared to 43.8% (95%CI: 43.4-44.2).

The inclusion of interaction terms between key predictor variables made no material difference to overall estimates of the stage distribution or the change over time. For example, the stage I/II

percentage in 2008-09 was estimated at 31.8% (95%CI: 31.1-32.6) with model (F) compared to 31.9% (95%CI: 31.3-32.5) with model (E).

Table 2. Comparison of estimates of early-stage colorectal cancer diagnosis, and change over time, based on different multiple imputation models, England, 2008-2013

Model specification	Stage I/II percentage (95% confidence interval)			Odds ratio for difference in stage I/II % between 2008-09 and 2012-13
	2008/09	2010/11	2013/12	
(A)	29.4 (28.7, 30.1)	38.8 (38.2, 39.3)	44.7 (44.3, 45.1)	1.9 (1.9, 2.0)
(B)	29.4 (28.8, 30.0)	38.8 (38.3, 39.3)	44.7 (44.3, 45.1)	1.9 (1.9, 2.0)
(C)	30.1 (29.4, 30.9)	38.1 (37.6, 38.6)	43.7 (43.2, 44.1)	1.80 (1.7, 1.9)
(D)	30.1 (29.4, 30.8)	38.1 (37.6, 38.6)	43.7 (43.2, 44.1)	1.80 (1.74 1.9)
(E)	31.9 (31.3, 32.5)	38.5 (38.1, 39.0)	43.8 (43.4, 44.2)	1.7 (1.6, 1.7)
(F)*	**31.8 (31.1, 32.6)**	**38.5 (38.0, 39.0)**	**43.8 (43.4, 44.2)**	**1.7 (1.6, 1.7)**
***Model (F) is considered *a priori* the most accurate model**				

Phase 2

Estimates from the imputation models where 40% of patients had had their stage re-assigned to missing were very similar to the true distribution of recorded stage, confirming that this MI approach works well under one plausible MAR mechanism (Table 3). Overall, 40.3% (95% CI: 40.2%, 40.7%) percent of these patients in 2008/09 were estimated to be diagnosed stage I or II compared to the true value of 40.5%, and estimates of the change during 2008-2013 were identical between the actual data and the imputed datasets to two decimal places.

Table 3. Comparison between the stage distribution of the 119,488 patients with stage recorded, and estimates of this distribution after stage was recoded to missing for 40% of the 119,488 patients: colorectal cancer in England 2008-2013.

	Stage I/II percentage (95% confidence interval)			**Odds ratio for 2008/09 to 2012/13 change from multilevel model***
	2008/09	**2010/11**	**2013/12**	
Actual stage distribution (N=119,488)	40.5%	31.0%	39.8%	1.82 (1.76, 1.88)
Multiple imputation estimates of stage distribution (95% confidence interval)	40.3% (40.0%, 40.7%)	30.9% (30.2%, 31.5%)	39.8% (39.2%, 40.3%)	1.82 (1.76, 1.88)
* Logistic regression model including parameters diagnosis period (2008-09, 2010-11, 2012-13) and random effects for patient's CCG of residence in each period.				

Phase 3

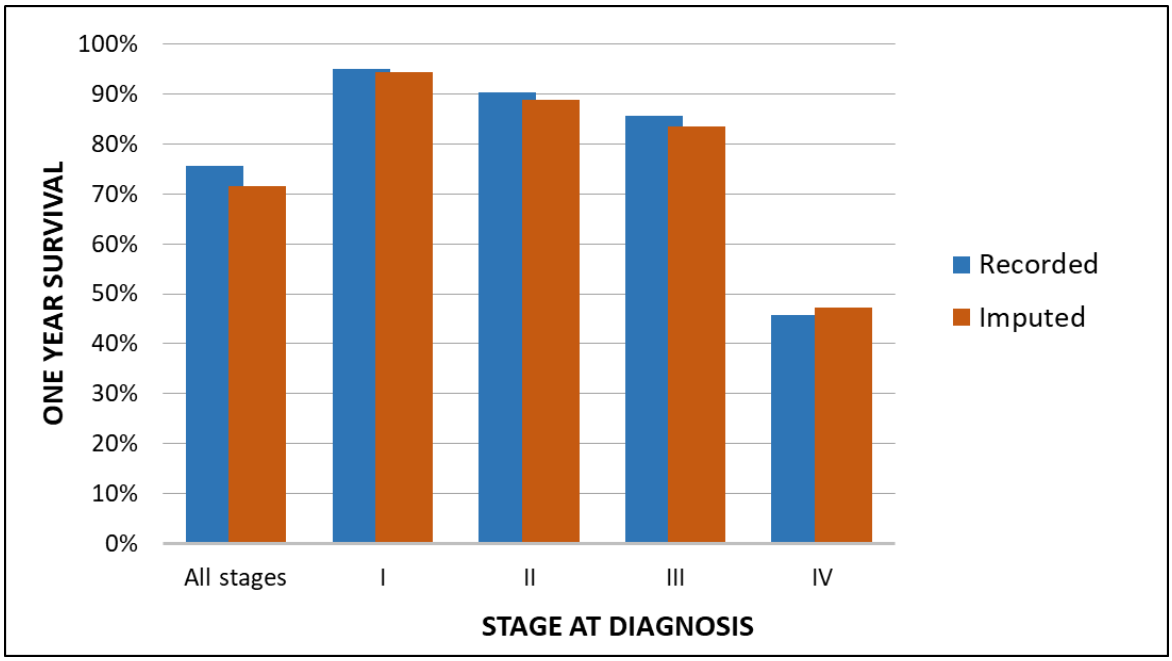
The pattern mixture sensitivity analysis indicated that substantial bias would have to present in the imputation to change the inferences about the early diagnosis changes during 2008-2013 we previously reported (Table 4). Considering only the 40% of patients whose stage was imputed, the improvement would be half as large if a net of 20% of those patients categorised as ‘late’ in the imputation model actually had ‘early’ stage disease. If 40% of the patients categorised as late had their stage incorrectly classified in this manner, then there would be no improvement.

Table 3. Sensitivity of multiple imputation early diagnosis estimates to breaches of the missing-at-random (MAR) assumption: effect of recoding different percentages of records with missing stage imputed III/IV to I/II.

	Overall stage I/II %	2008/09 stage I/II %	2010/11 stage I/II %	2012/13 Stage I/II %	Odds ratio for change in stage I/II diagnosis from 2008/09 to 2012/13 from multilevel model*
No recoding	38.2 (37.9, 38.5)	31.9 (31.4, 32.5)	38.5 (38.1, 39)	43.8 (43.4, 44.3)	1.66 (1.61, 1.72)
Percentage of records originally imputed as stage III/IV recoded to stage I/II:					
10%	40.7 (40.4, 41)	36.2 (35.6, 36.7)	41 (40.5, 41.5)	44.8 (44.4, 45.3)	1.43 (1.39, 1.48)
20%	43.3 (43, 43.6)	40.4 (39.9, 41)	43.5 (43, 44)	45.9 (45.4, 46.3)	1.25 (1.21, 1.29)
30%	45.9 (45.6, 46.1)	44.7 (44.1, 45.2)	46 (45.5, 46.4)	46.9 (46.5, 47.3)	1.09 (1.06, 1.12)
35%	47.1 (46.8, 47.4)	46.7 (46.2, 47.3)	47.2 (46.7, 47.7)	47.4 (47, 47.8)	1.03 (1.00, 1.06)
40%	48.4 (48.1, 48.7)	48.9 (48.4, 49.4)	48.4 (47.9, 48.9)	47.9 (47.5, 48.3)	0.96 (0.94, 0.99)
50%	51 (50.7, 51.2)	53.2 (52.7, 53.7)	50.9 (50.4, 51.4)	48.9 (48.5, 49.3)	0.84 (0.82, 0.87)
* Based on a logistic regression model with multiple imputation for missing data					

Figure 1 shows the one-year survival for the patients analysed, by whether stage at diagnosis was recorded I, II, III, or IV; or imputed (i.e. missing prior to the multiple imputation process) to I, II, III, or IV. Patients with imputed stage generally had slightly lower survival. For example, patients with recorded stage I had 95.1% (95% CI: 95.1%, 95.2%) survival at one year, compared to 94.3% (94.2%, 94.3%) for patients with imputed stage I (Figure 1, Appendix Table 1). This difference may be partly explained by other prognostic characteristics of patients without stage recorded. With respect to the accuracy of stage imputation, these data suggest late diagnosis is more likely to be under-estimated rather than over-estimated; that any bias is in the opposite direction to the one which would invalidate the conclusions of our previous analysis.

Figure 1. One year survival of colorectal cancer patients, overall and by stage at diagnosis according to whether it was recorded or imputed: England, 2008-2013



DISCUSSION

Key findings

We found a notable improvement in the accuracy of multiple imputation estimates of the stage distribution of colorectal cancer in England when patient's survival time was included in the imputation model. One year's survival was sufficient for accurate estimates when the missing stage percentage was 40%. By contrast, inclusion of interaction terms between diagnosis period and patient characteristics had almost no effect on imputed estimates of the stage distribution.

In a previous study we reported a very large increase in early-stage diagnosis of colorectal cancer during 2008-2013 using a model that included survival time (from 32% to 44%). Our sensitivity analyses indicate that finding is robust: 20% of the patients with imputed late-stage disease would have had to actually have had early-stage disease for the increase we report to be halved; 40% for no change at all. Patterns of survival by stage between patients with imputed and recorded stage indicate this is unlikely.

Strengths and limitations

A strength of this analysis is that we explored a wide range of auxiliary variables and model specifications to optimise the accuracy of our imputation model. We further confirmed the imputation model works well in one plausible scenario where stage was missing more frequently for patients in earlier periods and those with low survival, increasing confidence in this approach. The key limitation is that we could only compare different models to one *a priori* optimal model, and could not confirm that it itself is unbiased. However, using a pattern-mixture sensitivity analysis we showed that the risk that false conclusions were drawn in our previous analysis using this approach is very unlikely.

Aside from stage, other factors potentially leading to poor survival include pre-existing comorbidities; lack of access to treatment; grade; and serious adverse events at the time of cancer diagnosis. Our imputation accounted for these factors to a degree, but more specific variables may have improved it. We included a variable for major surgical treatment, but no data on receipt of radiotherapy or chemotherapy was available. We accounted for emergency presentation, which is a proxy for severe adverse events at time of diagnosis, but granular information on the initial diagnosis and treatment at point of emergency presentation was not included. Finally, we included morphologic cell type, which is associated with disease grade, but not grade itself, which is less well recorded

Comparison with other studies

Other studies have addressed the topic of stage imputation in epidemiological research, though none have previously evaluated early diagnosis trends. Barclay *et al* reported MI early-stage estimates which were consistently 1-2% lower than the complete case estimates for a group of common cancers in England [16], consistent with our findings. Public Health England have also used MI for missing stage when estimating survival by stage [17]. They imputed stage using cancer registry, sex, income deprivation, age, survival, and tumour metastatic status (yes or no). Metastatic status data was not included in our imputation. It is useful in cases where there is not enough information to assign the tumour a stage [11], but where a negative test for metastases excludes stage IV disease. Overall 5% of patients missing stage in our dataset met this criterion, indicating that although the information is valuable it is unlikely to change conclusions. For an analysis of socioeconomic inequalities in early-stage at diagnosis, using data for 10 common cancers in England in 2015, Barclay *et al* also employed a similar approach to that taken in this analysis, though their imputation model additionally included basis of diagnosis (microscopic, non-microscopic, or death certificate only), and survival time to one year only (parameterised as the cumulative hazard and two event indicators: one for 30 days and one for 365), but did not include routes to diagnosis (instead using a variable for whether the tumour was screening detected or not) or pre-existing comorbidities [18]. Other studies in the East of England have evaluated socioeconomic and age differences in stage, and employed multiple imputation as a sensitivity analysis, using a similar group of variables to this study and also including basis of diagnosis [19, 20]. Basis of diagnosis would likely have been a valuable addition to the imputation model in the present study, as microscopic diagnosis is expected to be more common for patients diagnosed at earlier stages, when the tumour is operable and histologic information can inform treatment decisions. A strength of the present study however is the inclusion of pre-existing comorbidities, as adjustment for these is expected to improve the performance as survival as a proxy for stage.

Falcaro *et al* performed a re-sampling analysis to impute colorectal cancer stage for patients diagnosed in England between 1996 and 2006, to calculate stage-specific net survival [21]. Setting 50% of recorded stage data to missing under one MAR scenario based on survival, age, and deprivation, they then imputed stage using data on these characteristics. MI was found to almost completely eliminate bias in the survival estimates. A further study on colorectal cancer stage in England solicited opinions from six clinical experts about the values the missing stage takes, then

combined information from these and the predictive distribution from the MAR imputation model [22]. Results integrating the expert's opinions were broadly similar to the initial results.

One study in Germany compared multinomial regression, predictive mean matching, the random forest classification algorithm, and proportional sampling for stage imputation for malignant melanoma and breast cancer [23]. Multinomial regression – the approach used in this present study – produced the most accurate imputation. Another study by Falcaro *et al* found that estimates of net survival by stage were sensitive to the handling of missing stage information (in this case stage being a covariable rather than the outcome), and that multinomial logistic imputation models were the best means of minimising bias [24].

A study on prostate cancer in New South Wales used MI to handle missing stage in cancer registration records, and compared the results to the patient's actual stage which had been ascertained in clinical notes captured as part of a population-based cohort study [25]. They compared one imputation model including age, geography, deprivation, survival status and survival time, to another which also included treatment within 6 months (from different surgical options, radiotherapy, and watchful waiting). Both MI models were found to return minimally biased stage estimates compared to the actual ascertained stage. He *et al* describe a similar approach, where information on a subsample of patients gained from a clinician survey was combined with multiple imputation estimates of adjuvant treatment receipt using Bayes theorem [26], In that study, the California Cancer Registry was found to underreport adjuvant treatment when compared to clinician's records, but bias from this was minimised by incorporating the survey results in the imputation process.

Conclusion

Early diagnosis remains a priority of cancer research in the UK [10, 27]. The 2019 NHS Long Term Plan has set a target that by 2028 the proportion of cancers diagnosed at stages I or II in England should rise to 75% [10]. Robust population-based data on early diagnosis is essential to monitoring national progress against this target, and geographic inequalities within England. However, the challenge from missing data on early diagnosis needs to be addressed convincingly and not ignored. Though completeness of stage recording improved considerably to 2018 [1], a significant minority of patients still lack a recorded stage, and they have poorer outcomes [2].

Our study and others demonstrate that multiple imputation is the most appropriate approach to assess progress in early diagnosis in the whole population, as it allows the wealth of contextual information on patients whose stage was not recorded to be used to accurately impute it. We have shown that, for colorectal cancer in England, information on one year's survival time from diagnosis is needed to return robust estimates. Other key variables needed for imputation are those which are associated with prognosis, such as receipt of treatment and route to diagnosis. With this approach, a more robust assessment of progress in early diagnosis in the whole population can be achieved. The same methods would also confer benefits in other countries where the recording of stage is poor, but where other contextual information on cancer patients is available.

FIGURE TITLES

Figure 1. One year survival of colorectal cancer patients by stage at diagnosis (recorded or imputed), England, 2008-2013

FOOTNOTES

Contributors PM and LW designed the study. PM did the analysis and drafted the manuscript under supervision from LW. Both authors contributed to the interpretation of the study results and preparing the manuscript, and approved the final draft for submission.

Funding PM received funding from a Cancer Research UK (CRUK) Early Diagnosis Policy Research grant to the Cancer Policy Programme at the London School of Hygiene & Tropical Medicine (award number C7923/A18348), and LW received funding from a Cancer Research UK Population Research Postdoctoral Fellowship (award number C23409/A11415).

Competing interests None.

Patient consent for publication Not required.

Ethics approval This research was approved by the Observational Research Ethics Committee at the London School of Hygiene & Tropical Medicine.

Data sharing statement: The data were obtained from Public Health England. The authors do not own them and hence are not permitted to share them in their original form.

REFERENCES

1 Public Health England. CancerData dashboard. 2019.

2 Muller P, Walters S, Coleman MP*, et al.* Which indicators of early cancer diagnosis from population-based data sources are associated with short-term mortality and survival? *Cancer Epidemiol* 2018;**56**:161-70.

3 Logan RFA, Patnick J, Nickerson C*, et al.* Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. *Gut* 2011;**61**:1439-46.

4 Department of Health. Overview of the the Health and Social Care Act fact sheet. London, UK 2012.

5 Maringe C, Spicer J, Morris M*, et al.* The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study. *The Lancet Oncology* 2020;**21**:1023-34.

6 Sterne JAC, White IR, Carlin JB*, et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;**338**.

7 Rubin DB. *Multiple imputation for nonresponse in surveys*. New York: John Wiley and Sons 1987.

8 Muller P, Woods L, Walters S. Temporal and geographic changes in stage at diagnosis in England during 2008–2013: A population-based study of colorectal, lung and ovarian cancers. *Cancer Epidemiology* 2020;**67**:101743.

9 Carpenter JR, Kenward MG. Missing data in randomised controlled trials: a practical guide. Birmingham, UK: Health Technology Assessment Methodology Programme 2008.

10 NHS England. The NHS Long Term Plan. 2019.

11 Benitez Majano S, Fowler H, Maringe C*, et al.* Deriving stage at diagnosis from multiple population-based sources: colorectal and lung cancer in England. *Br J Cancer* 2016;**115**:391-400.

12 Elliss-Brookes L, McPhail S, Ives A*, et al.* Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets. *Br J Cancer* 2012;**107**:1220-6.

13 Maringe C, Fowler H, Rachet B*, et al.* Reproducibility, reliability and validity of population-based administrative health data for the assessment of cancer non-related comorbidities. *PLoS One* 2017;**12**:e0172814.

14 Quartagno M, Carpenter J. jomo: A package for Multilevel Joint Modelling Multiple Imputation. 2017.

15 StataCorp. meqrlogit - Multilevel mixed-effects logistic regression (QR decomposition). 2013.

16 Barclay ME, Lyratzopoulos G, Greenberg DC*, et al.* Missing data and chance variation in public reporting of cancer stage at diagnosis: Cross-sectional analysis of population-based data in England. *Cancer Epidemiology* 2018;**52**:28-42.

17 National Cancer Intelligence Network. Cancer survival in England by stage. London: Public Health England 2014.

18 Barclay ME, Abel GA, Greenberg DC, et al. Socio-demographic variation in stage at diagnosis of breast, bladder, colon, endometrial, lung, melanoma, prostate, rectal, renal and ovarian cancer in England and its population impact. British Journal of Cancer 2021;124:1320-9.

19 Lyratzopoulos G, Abel GA, Brown CH, et al. Socio-demographic inequalities in stage of cancer diagnosis: evidence from patients with female breast, lung, colon, rectal, prostate, renal, bladder, melanoma, ovarian and endometrial cancer. Ann Oncol 2013;24:843-50.

20 Lyratzopoulos G, Abel GA, Barbiere JM, et al. Variation in advanced stage at diagnosis of lung and female breast cancer in an English region 2006-2009. Br J Cancer 2012;106:1068-75.

21 Falcaro M, Carpenter JR. Correcting bias due to missing stage data in the non-parametric estimation of stage-specific net survival for colorectal cancer using multiple imputation. *Cancer Epidemiology* 2017;**48**:16-21.

22 Smuk M, Carpenter JR, Morris TP. What impact do assumptions about missing data have on conclusions? A practical sensitivity analysis for a cancer survival registry. *BMC Med Res Methodol* 2017;**17**:21.

23 Eisemann N, Waldmann A, Katalinic A. Imputation of missing values of tumour stage in population-based cancer registration. *BMC Medical Research Methodology* 2011;**11**:129.

24 Falcaro M, Nur U, Rachet B, et al. Estimating excess hazard ratios and net survival when covariate data are missing: strategies for multiple imputation. Epidemiology 2015;26:421-8.

25 Luo Q, Egger S, Yu XQ*, et al.* Validity of using multiple imputation for "unknown" stage at diagnosis in population-based cancer registry data. *PLOS ONE* 2017;**12**:e0180033.

26 He Y, Yucel R, Zaslavsky AM. Misreporting, Missing Data, and Multiple Imputation: Improving Accuracy of Cancer Registry Databases. Chance (New York, NY) 2008;21:55-8.

27 Cancer Research UK. Early detection and diagnosis of cancer: a roadmap to the future. 2020.

12.2 Supplementary appendices

Appendix 1: Stata code recode stage to missing under a missing at random (MAR) mechanism

```
use "CRC_2008_2013_FOR_MI_R.dta", clear

keep if stage !=.

gen rand = runiform() /* random uniform variable, values in range [0,1] */

codebook H if bin_time08 == 1

codebook H if bin_time12 == 1

codebook diagmdy /* average diagnosis date is 18831, minimum 17532 */

gen centr_diagmdy = diagmdy - 18831

replace centr_diagmdy = (centr_diagmdy/1299)+1 /* range [0,1.6], bigger=later */

replace rand =rand - (centr_diagmdy)/2 /* rand now lower if diagnosed later */

replace rand = rand*1.4 if H < 0.15 & d == 1 /* rand now higher if survival low */

replace rand = rand*1.2 if H > 0.14999 & H <0.35 & d == 1 /* as above */

sort rand

gen n=_n

codebook n

display 0.6*119488 /* need to set 47795 to missing, going to bias selection towards higher values of rand (i.e. on average the patients who were diagnosed earlier, or with lower survival) but with some coverage of all values of rand */

replace stage =. if n > 104488 & n < 114488 /* 10k */

replace stage =. if n > 85488 & n < 98488  /* 13k */
```

```
replace stage =. if n > 63488 & n < 75488 /* 12k */

replace stage =. if n > 50488 & n < 55488 /* 5k */

replace stage =. if n > 25488 & n < 30488 /* 5k */

replace stage =. if n > 5000 & n < 7795 /* 2795 */

saveold "CRC_2008_2013_FOR_MI_R_60only.dta", version(12) replace
```

Appendix Table 1. One year survival by stage at diagnosis (recorded or imputed), colorectal cancer, patients diagnosed in 2008-2013

12-month survival of patients with stage recorded (N=119,488), % (95% CI)		12-month survival of patients without stage recorded (N=77,023), % (95% CI)	
Recorded: overall	75.6 (75.5, 75.6)	Imputed: overall	71.6 (71.6, 71.7)
I	95.1 (95.1, 95.2)	I	94.3 (94.2, 94.3)
II	90.4 (90.4, 90.5)	II	88.9 (88.8, 88.9)
III	85.6 (85.6, 85.7)	III	83.5 (83.4, 83.6)
IV	45.7 (45.7, 45.8)	IV	47.3 (47.2, 47.4)

12.3 Implications for monitoring of early cancer diagnosis and future research

Improving the timeliness of reporting of surveillance data is a common priority of clinical audit and health surveillance programmes in England. In response to the COVID-19 pandemic, Public Health England commenced publication of rapid cancer registration and treatment data, with only four months delay between data collection and reporting [181]. The dashboard authors caution this data is best-suited for service improvement and planning compared to registrations which have gone through quality checks and cross referenced against other data sources, stating “it [the registration data] is poorly suited for epidemiological research due to limitations in the data quality and completeness”. In particular, quality checks reported alongside the data highlight that rapid stage reporting based on it will be inaccurate.

The present study suggests that data on one year’s survival from time of diagnosis is needed for an optimal imputation approach to handle missing stage data for one common cancer. However, further research is needed to confirm this finding and test if it holds for other cancers. Furthermore, due to a

maximum follow up of December 2014 for all patients analysed, imputations using one year's follow up and up to six years follow up for patients diagnosed in 2012-13, is practically the difference between one year's follow up and up to three years follow up. Our finding of very similar early-stage percentages between these two follow up periods needs to be considered in this context. A further analysis, with maximum follow up to 2019, would be helpful to confirm that these differences remain small in a comparison between one year follow up and a full six years for each patient.

Together with the limitations of rapidly available stage data, the findings suggest there is a need for epidemiological analyses of stage trends based on fully cleaned stage data, with multiple imputation to handle missing values. Such analyses will provide more robust feedback on the impact of different interventions and events, but only some time (i.e. at least over a year) after they occurred. These may be complemented by rapid processing and reporting of the immediately available data, for interim assessment before a full evaluation is possible.

Whilst the results of this research project broadly support the approach taken in the substantive analysis of the thesis, some areas which could have been improved on emerge through examination of other studies in this area. Other analysts have included an indicator variable for death within 30 days of diagnosis and basis of diagnosis in imputation models, which could both potentially have provided more accurate imputation results. The comorbidity variable I used is based on diseases recorded between 6 and 60 months prior to diagnosis. One possible extension would be to also include diseases recorded around the time of diagnosis: adjustment for these would allow the imputation to take account of incidental cancer diagnoses whilst the patient was being treated for another condition, which in some cases may have been more important to the patient's short-term survival than the cancer.

13. Discussion

13.1 Summary of thesis findings

Increasing early diagnosis is a key objective of cancer policy in England. Early diagnosis statistics allow evaluation of progress, and can provide a powerful stimulus for continuous improvement. I identified several potential challenges to generating robust statistics, however. These included confounding from case mix, sparse data for comparisons between local areas, and missing data on stage at diagnosis. I aimed to address these issues, and demonstrate optimal methods for monitoring trends and geographic inequalities in early diagnosis for three common cancers in England during 2008-2013. I also aimed to produce recommendations for monitoring up to 2021 and beyond.

A modelling strategy to evaluate trends and inequalities whilst accounting for missing stage data and case-mix differences was identified. Deploying this strategy, I found evidence for an increase in the percentage of patients diagnosed at stages I or II for each of colorectal, non-small lung, and ovarian cancer during 2008-2013. Two key findings were of a large, step-change increase in early-stage diagnosis of colorectal cancer, from 32% of patients in 2008-09 to 44% 2012-13, and for non-small cell lung cancer, from 19% of patients in 2008-09 to 25% in 2012-13. For colorectal and ovarian cancer there was also evidence of a reduction in geographic inequalities in early diagnosis, and a shift from stage IV to stage III disease amongst patients who were diagnosed at these later stages.

I found that adjustment for case-mix differences made no material difference to estimates of trends or geographic inequalities for the cancers I analysed during 2008-2013. Estimates were more sensitive to the handling of missing stage data, with some overestimation of early-stage diagnosis in complete case analyses. Application of appropriate methods is needed to handle both potential sources of bias in future analyses.

With respect to handling of missing stage data, there was evidence of greater disparities in survival and routes to diagnosis between patients with stage recorded and patients without stage recorded in 2012-13 than in 2008-09. The patients missing stage had comparatively worse outcomes in 2012-13, by which time recording of stage had increased. This has important implications for future monitoring. Completeness of stage recording was greatly improved after 2013, but if precisely the patients with the worst outcomes are the ones with missing data, then material bias may be present. If these patients with missing stage are not accounted for appropriately, monitoring may overestimate early diagnosis.

Sensitivity analyses indicate that the multiple imputation model was appropriately specified for colorectal cancer. Additionally, in a simulation where stage values were recoded to missing at random for 40% of patients with stage recorded, and then accounted for using multiple imputation, accurate estimates of trends were returned. The sensitivity tests also confirmed that overall conclusions about trends for that cancer were robust to residual bias.

The sensitivity analyses also indicated that data on one year's survival is ideal to generate robust multiple imputation estimates for colorectal cancer. Together with the limitations of rapidly-reported stage data described by reports of the Cancer Registry [181], this finding implies that, at least for some cancers, it is not currently possible to do an assessment of stage trends that is both rapid and robust in England. The immediately available stage data may however be used, alongside data on routes to diagnosis, for interim assessments, which can then be followed by less timely but more robust analyses of overall trends.

13.2 Key results supporting the thesis findings

13.2.1 Indicators of early diagnosis

I evaluated three indicators which could potentially be used to monitor early diagnosis in the whole population in England: stage, emergency presentation, and interval from symptoms to diagnosis. Stage was the only indicator which was associated with patient survival, and was also a causal determinant. It was therefore chosen as the outcome for the substantive analysis of trends. Interval from symptoms to diagnosis was not associated with survival in population-based data sources, likely due to confounding from disease severity. Emergency presentation is strongly associated with survival, but as an administrative classifier it has a more complex interpretation than stage, and its prognostic value may change over time for different cancers.

An analysis of survival by stage provided insights for the substantive analysis. Each stage is associated with different long-term survival, and so analyses of trends should measure changes in the distribution of all stages I-IV, alongside or instead of measures which dichotomise stage into "early" and "late". Patients without disease stage recorded in their cancer registration (20-40% of all patients during 2008-2013) had poorer survival than patients with stage recorded. For example, overall, one-year net survival for patients diagnosed with NSCLC aged 60-79 was 35.0% compared to just 24.4% for those without a recorded stage. This finding indicated that bias could be introduced by excluding the patients with missing stage (e.g. a complete case analysis), and that these patients should be accounted for in estimates of trends.

13.2.2 On case-mix adjustment

I created a conceptual framework to map the relationship between patient factors, system factors, and early diagnosis. This indicated that age, sex, comorbidity, and tumour morphology and topography should be adjusted for to generate statistics which are specific to health services performance. Including these items as a covariables in a generalised linear model was identified as one way to achieve this.

13.2.3 Techniques for evaluating geographic inequalities

An initial funnel plot analysis revealed material between-CCG variation in early diagnosis during 2008-2013, beyond what would be expected due to the random play of chance. However, the funnel plot

analysis also did not return evidence that any one CCG or group of CCGs had extreme stand-out results against the overall pattern of variation. Because of sparse data at CCG level and the nature of the variation observed, I decided to use a modelling approach to evaluate whether total between-CCG variation increased or decreased during 2008-2013. Multilevel generalised linear models, with parameters for overall between-CCG variation in different time periods within 2008-2013, were identified as one viable approach.

13.2.4 Handling of missing data

Multiple imputation was identified as the ideal approach to reduce bias from missing stage data in estimates of stage trends and geographic inequalities. The patients missing a record of stage had other information available on demographics, route to diagnosis, treatment, and survival time. Multiple imputation allowed this information to be used to impute stage where it was not recorded, resulting in more plausible assumptions about the stage distribution of these patients than a complete case analysis. An assessment of the impact of including different parameters in imputation models showed that, in an analysis of colorectal cancer trends, interaction parameters between calendar time and other variables had minimal effect. Parameters for patient's survival time from diagnosis were important for accurate stage imputation. Though longer-term data on survival was preferable, I found when missing stage was low (<15%), information on survival to one year was adequate to minimise bias. More generally, stage imputation models should contain survival time and all other variables which may strongly predict survival (e.g. comorbidities, treatment, morphology and grade if recorded).

13.2.5 Estimates of changes in early diagnosis during 2008-2013

Multiple imputation estimates provided evidence of significant improvements in early diagnosis of the three cancers I considered during 2008-2013. The percentage of patients diagnosed at stage I or II increased from 32% to 44% for colorectal cancer; from 19% to 25% for NSCLC; and from 28% to 31% for ovarian cancer. There was also a favourable stage shift from IV to III for colorectal and ovarian cancers, and an increase in diagnoses at stage I for all cancers, and good evidence that geographic inequalities in early diagnosis reduced by 2013 for colorectal and ovarian cancer. Referral guidelines and waiting times targets may have contributed to these reductions by standardising practices in different CCG areas.

The large early diagnosis increase for colorectal cancer occurred concurrently with the introduction and then expansion of the FOBT screening programme, and increasing availability of flexible sigmoidoscopy. The percentage of colorectal cancer patients diagnosed through screening rose from 6% to 10% during 2008-2013, the percentage diagnosed on the urgent GP referral pathway rose from 26% to 30%, while diagnoses through emergency presentation fell. These changes provide evidence of more patients presenting to health services earlier in the disease progression, and then having shorter waits for diagnosis. These are likely to have played a significant role in the improvement of early diagnosis. For NSCLC there was no equivalent screening programme in this period, but symptom awareness campaigns and improved treatment outcomes at this time may have led to a change in attitudes with respect to early diagnosis amongst patients and clinicians [182].

In general, estimates of trends during 2008-2013 did not vary substantially between complete case and multiple imputation analyses. The complete-case improvement in early diagnosis for colorectal cancer was 31% to 45% (compared to 32% to 44% with multiple imputation); for NSCLC it was 20% to 26% (compared to 19% to 25%); for ovarian cancer it was 33% to 33% (compared to 28% to 31%). The lower estimated early-stage diagnosis percentage with multiple imputation is consistent with the poor survival of patients missing stage, and confirms that complete case analyses typically over-estimate early diagnosis. Furthermore, for ovarian cancer, a different conclusion about trends was reached with the complete case analysis (no change) compared to multiple imputation analysis (small improvement). This finding highlights that multiple imputation should be done either as a primary or sensitivity analysis as, if not addressed, bias from missing data may lead to different conclusions about trends than would otherwise be reached.

One sensitivity analysis, conducted after the substantive analysis, showed that considerable over-estimation of late-stage diagnosis for the patients missing stage would have had to be present for the conclusions about the colorectal stage trends during 2008-2013 to be invalid.

13.3 Comparison with other studies

At the time this thesis was started in 2016, Public Health England's online *CancerData* dashboard was the only source of information on stage trends in England [5]. As of February 2021, the dashboard continues to report the (complete case) percentage of patients diagnosed at stages I or II from a basket of 10 common cancers. New pages have also been added to the website, which include information on counts of patients by stage at diagnosis (I, II, III, IV, or missing) from 2013 to 2018, and data quality statistics [57]. Changes in the complete-case early-stage percentage can be estimated from these online data. Where similar definitions are used, the results are broadly similar to mine. For colorectal cancer, the reported stage I/II percentage remained flat during 2013-2018 at approximately 44.5%. For lung cancer (including small cell cancers) the stage I/II% increased from 24% in 2013 to 29% by 2018. *CancerData* uses a different definition of ovarian cancer than the one in this thesis, excluding malignant neoplasms of other specified female genital organs (C57.7) but including primary peritoneal cancers (C48). Using that definition, the stage I/II percentage remained flat at approximately 40% in both 2013 and 2018. The dashboard shows that the missing stage percentage dropped substantially during 2013-2018 for each these cancers.

As discussed in detail in Chapter 12, other studies have used multiple imputation to estimate stage distributions, though none analysed national time trends in early-stage diagnosis. The methodology for imputation, and results from it, in other studies are broadly similar to mine. Barclay *et al* (2019) estimated the early-stage percentage for a group of common cancers in England in 2013, reporting estimates consistently 1-2% lower than complete case results [44], similar to my analysis. Other studies reported that multiple imputation minimised bias in estimates of stage-specific survival; that the multinomial regression approach performs better than other approaches; and one reported that, in

Australia, multiple imputation estimates based on registry records provide good estimates of prostate cancer stage when compared to granular patient notes [183-185].

The COVID-19 pandemic has greatly disrupted NHS services, and it is likely to significantly impact cancer policy and health policy in future. Early studies have sought to assess its impact on early diagnosis (though none have yet estimated stage trends). From March 2020 cancer screening appointments were delayed, routine hospital appointments were cancelled, and patients were initially discouraged from consulting their GP for anything except emergencies. Maringe *et al* estimated the impact of diagnostic delays due to COVID on survival [186]. Survival by route to diagnosis data was used to estimate the impact if all the patients diagnosed through screening or routine referral pathways (halted) were instead diagnosed following emergency presentation or urgent GP referral (still operating). The study estimated there would 3,000 excess avoidable deaths due to the service interruption if it continued for 12 months. This finding underlined the importance of re-establishing patient flows through the health system where possible. As route to diagnosis is not (by itself) a causal determinant of survival, and may not have a consistent relationship with survival as other factors changed in the crisis, the results were approximate – though highly successful in rapidly drawing attention to the hazard to patients from disruption to normal pathways through the health system. One alternative to the approach taken in the study by Maringe *et al* would be to model the impact of the service changes, including diagnostic delays, on stage directly, and then estimate survival by stage using these. However, due to the paucity of reliable data on the causal effects of diagnostic delays, such a study would return very approximate results.

13.4 Strengths and limitations

A strength of this thesis was the wide range of data available. The analyses conducted included data from cancer registrations, site-specific clinical audits, routes to diagnosis, and comorbidities (derived from HES inpatient records). Data on survival to at least one year after diagnosis was also available for all patients. The availability of this comprehensive data improved the plausibility of the imputation of missing stage information based on other data captured on the patients.

Another strength was the thorough consideration of methodological issues. Numerous potential pitfalls were avoided: bias from missing data was reduced through multiple imputation; and sparse data did not compromise the analysis of geographic inequalities. Sensitivity analyses confirmed the conclusions from multiple imputation were robust to some mis-estimation of stage. Stage has a direct causal effect on survival, and so is a readily interpretable indicator. By contrast, as discussed in Section 13.3, routes to diagnosis may have a more complex interpretation, with the prognostic value of different routes changing over time. Survival differences by route may also be more affected by lead time bias: the phenomena by which the patient's date of diagnosis is brought forward, but their date of death remains unchanged [187]. Lead time bias could potentially inflate the survival benefit of diagnosis through the two-week wait or screening pathways compared to other routes, though long-term and conditional survival are less sensitive to this.

A key intention of the thesis was to return statistics which could provide feedback on the success of early diagnosis interventions and the impact from organisational change; to provide a "health check" of the health service's performance. Unfortunately, data beyond 2013 were not available for analysis. These would have been used to assess the functioning of the health service after 2013, following the NHS reorganisation and during ongoing financial austerity. Complete-case statistics from 2013-2018 on the *CancerData* dashboard do indicate that early diagnosis improvements stalled after 2013 for colorectal and ovarian cancer, but continued for lung cancer. Further analyses using multiple imputation are now necessary to confirm this.

At the point of publication in *Cancer Epidemiology* in 2020 the data from the substantive analysis were seven years old, reducing the potential for decisions to be made based on them. However, those results do speak to the success of interventions during 2008-2013, in particular new screening and diagnostic technologies. Other countries with persistently low colorectal cancer and lung cancer survival could look to implement the policies for screening and diagnosis which were actioned in England, on the strength of this evidence of their success up to 2013. To inform current cancer policy in England, the thesis primarily provides a methodological basis for future monitoring, in particular handling of missing stage data.

Another limitation is that generally only ecological data on different interventions to raise early diagnosis was available, and the effect of each intervention on early diagnosis for a specific patient was not known. Information on routes to diagnosis did allow assessment of increases in diagnoses following urgent GP referral (and for colorectal cancer, through screening), however.

There is likely to be some residual bias from missing stage data with multiple imputation, of an unknown but assumed small degree. A future granular investigation of HES records to better understand the characteristics of patients missing stage, in collaboration with clinicians familiar with both the data entry systems and typical patient pathways, may yield insights to improve the imputation and better understand the circumstances of patients whose stage is not ascertained or not recorded for non-administrative reasons.

A final limitation is that the thesis conclusions are based on an analysis of three cancers, and may not be fully generalisable to other malignancies. For example, one year's survival was found to be sufficient for accurate stage imputation for colorectal cancer, but longer (or much shorter) survival data may be needed for other malignancies.

13.5 Implications for future research and policy

The findings of this thesis have several immediate applications. The COVID-19 pandemic greatly disrupted the normal functioning of health services in England from March 2020. In 2021 data will be available to conduct a robust evaluation of the impact of the pandemic on stage trends, using multiple imputation with survival time up to one year to minimise bias from missing data. If trends indicate a

negative effect on early diagnosis, the results can stimulate measures to ensure normal activity on the patient pathways reaches pre-pandemic levels, and also inform preparations for future emergencies.

Complete-case analyses indicate that progress in early-stage diagnosis stalled for colorectal cancer and ovarian cancer after 2013, but continued to increase for lung cancer. Multiple imputation stage trend estimates are needed to confirm these results. National stage trends are also the ideal means of evaluating the full impact of cancer-specific early diagnosis interventions in future. For example, a screening programme may be highly effective in a trial setting, but its ultimate effect at population level will depend on it being equally effective outside that setting; on participation from high-risk groups in the population; and on the functioning of patient pathways into and out of the screening programme. Population-based analyses of stage trends as performed in this thesis can confirm effectiveness, or highlight that improvements are needed. Where interventions are limited to a specific age group, early diagnosis trends can be compared between that group and others to try and isolate a distinct effect. Concurrent analyses of survival trends can complement evaluations of early diagnosis, though with the caveat mentioned in section 13.4 that lead-time play may play a role in any observed changes, particularly in the case of screening programmes.

One relevant finding of the thesis for future evaluations was that, even for common cancers, comparisons between individual CCG territories and the national average are underpowered even when several years of data were included. Pilot schemes of early diagnosis interventions being considered for national rollout should be either done a larger scale than CCGs (i.e. at a regional level) or with the understanding that they may be underpowered to detect improvements in early-stage diagnosis of a similar order of magnitude to the differences between best and average-performing CCGs. One potential analysis approach for an intervention that is already planned for national rollout is a stepped wedge cluster-randomised trial design [188], where the national rollout is organised as a trial where all areas start without the intervention and all are gradually allocated to it. This design can test the value of an intervention with minimal delay, and with higher power and reduced bias compared to a pilot scheme and a post-hoc ecological analysis.

The percentage of all cancer registrations without stage recorded dropped to <15% by 2018, and analysts may decide to ignore the missing data (i.e. conduct a complete case analysis) on the basis that the overall percentage is now small, and hence arising bias should be small. My findings indicate this approach would be incorrect for the cancers analysed in this thesis. The patients missing stage had poorer outcomes compared to patients with stage recorded, and the gap in outcomes widened during 2008-2013. This indicates that, as administrative practices have improved, missing stage increasingly identifies those patients with poorest outcomes.

13.5 Conclusion

Monitoring of early cancer diagnosis has the potential to play a key role in improving cancer services and outcomes in England, and internationally. Early diagnosis statistics can be used to understand

inequalities, evaluate interventions, and hold commissioners and policymakers to account for service performance, as cancer survival statistics have done in recent decades. Robust analyses of trends are challenging. However, this thesis demonstrates appropriate methods to conduct them, which minimise bias from missing data and case mix. Improvements in stage data completeness from 2013 should further reduce bias and uncertainty in assessments of trends. However, a key finding of this thesis is that the composition of patients missing stage has changed as stage completeness has improved, and they have considerably poorer outcomes than patients with stage recorded. Future application of the approaches used in this thesis will ensure that all patients with a cancer registration are counted in monitoring.

14. References

[1] Public Health England, Stage at Diagnosis 2012-2014 and One-year Cancer Survival in England, 2016.

[2] Department of Health, Cancer Reform Strategy, London, 2007.

[3] Independent Cancer Taskforce, Achieving World-Class Cancer Outcomes: A Strategy for England 2015-2020, London, 2016.

[4] G. Stoye, B. Zaranko, UK Health Spending, Institute for Fiscal Studies, 2019.

[5] Public Health England, CancerData Dashboard https://www.cancerdata.nhs.uk/dashboard#?tab=Overview, 2019 (accessed 25 April 2021).

[6] LaingBuisson, Health Cover UK Market Report, 2015.

[7] Office for National Statistics, UK Health Accounts, 2014.

[8] National Health Service and Community Care Act, The Stationery Office, London, 1990.

[9] The King's Fund, Clinical commissioning: what can we learn from previous commissioning models? https://www.kingsfund.org.uk/projects/nhs-white-paper/gp-commissioning, 2012 (accessed 25 April 21).

[10] BBC News, The changing role of health authorities http://news.bbc.co.uk/1/hi/health/background_briefings/your_nhs/93713.stm, 1999 (accessed 25 Apr 2021).

[11] Department of Health, Overview of the Health and Social Care Act Fact Sheet, London, 2012.

[12] NHS Clinical Commissioners, About CCGs. https://www.nhscc.org/ccgs/, 2021 (accessed 25 April 2021).

[13] Department of Health and Social Care Annual Report and Accounts 2018-19, Stationery Office, London, 2019.

[14] C. Ham, B. Baird, S. Gregory, J. Jabbal, H. Alderwick, The NHS under the coalition government Part One: NHS reform, King's Fund, London, 2015.

[15] J. Hammond, T. Mason, M. Sutton, *et al.*, Exploring the impacts of the 2012 Health and Social Care Act reforms to commissioning on clinical activity in the English NHS: a mixed methods study of cervical screening, BMJ Open 9(4) (2019) e024156. https://dx.doi.org/10.1136%2Fbmjopen-2018-024156.

[16] The King's Fund, Sustainability and Transformation Plans (STPs) explained. https://www.kingsfund.org.uk/topics/integrated-care/sustainability-transformation-plans-explained, 2017 (accessed 25 April 21).

[17] S. Neville, NHS calls for end to rules forcing services to be tendered, Financial Times. https://www.ft.com/content/3d657652-3b69-11e9-b856-5404d3811663, 2019 (accessed 25 April 21).

[18] C. Allemani, H.K. Weir, H. Carreira, *et al.,* Global surveillance of cancer survival 1995-2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2), Lancet 385(9972) (2014) 977-1010.

[19] Expert Advisory Group on Cancer, A Policy Framework for Commissioning Cancer Services, Department of Health, Stationery Office, London, 1995.

[20] C.N. Prabhakar, K.M. Fong, M.D. Peake, D.C. Lam, D.J. Barnes, The effectiveness of lung cancer MDT and the role of respiratory physicians, Respirology 20(6) (2015) 884-8.

[21] House of Commons Committee of Public Accounts, Delivering the cancer reform strategy. HC667, Session 2010-11, The Stationery Office, London, 2011, 1-22.

[22] Department of Health, The NHS Cancer Plan, London, 2000.

[23] R. Roope, New NICE GP guidelines have huge ambition and potential. https://scienceblog.cancerresearchuk.org/2015/06/23/new-nice-gp-guidelines-have-huge-ambition-and-potential/?_ga=2.101901505.1789612612.1564331422-821341162.1564331422 2015 (accessed 25 April 21).

[24] Department of Health, Improving Outcomes: A Strategy for Cancer, London, 2011.

[25] S. Duffy, K. Fenton, Be Clear on Cancer symptom awareness campaigns, London, 2014.

[26] National Health Service, NHS Five year Forward View, 2014.

[27] I. Blunt, N. Edwards, L. Merry, What's behind the A&E 'crisis'?, Nuffield Trust, London, 2015.

[28] B. Baird, A. Charles, M. Honeyman, D. Maguire, P. Das, Understanding pressures in general practice, The King's Fund, 2016.

[29] National Cancer Intelligence Network, Routes to Diagnosis: investigating the different pathways for cancer referrals in England for Teenagers and Young Adults, National Cancer Intelligence Network, London, 2013.

[30] National Institute for Health and Care Excellence (NICE), Suspected cancer: recognition and referral, 2015.

[31] Macmillan Cancer Support, The role of cancer networks in the new NHS. http://www.macmillan.org.uk/documents/getinvolved/campaigns/theroleofcancernetworksinthenewnhs.pdf, 2012 (accessed 25 April 21).

[32] NHS England, About the Cancer Vanguard. http://cancervanguard.nhs.uk/about/, 2021 (accessed 25 April 21).

[33] NHS England, Achieving World-Class Cancer Outcomes: Taking the strategy forward, 2016.

[34] Public Health England, What Does the NHS Breast Screening Programme Do?. http://www.cancerscreening.nhs.uk/breastscreen/screening-programme.html, 2015 (accessed 25 April 21)

[35] National Health Service, Bowel Cancer Screening. https://www.nhs.uk/conditions/bowel-cancer-screening/, 2015 (accessed 25 April 21).

[36] J. Gubb, O. Meller-Herbert, Markets in health care: the theory behind the policy, CIVITAS: Institute for the Study of Civil Society, London, 2009.

[37] C. Ham, Reforming the NHS from within: Beyond hierarchy, inspection, and markets, The King's Fund, 2014.

[38] J. Gubb, Have targets done more harm than good in the English NHS? Yes, Bmj 338 (2009) a3130.

[39] G. Bevan, Have targets done more harm than good in the English NHS? No, Bmj 338 (2009) a3129.

[40] G. Bevan, C. Hood, What's measured is what matters: targets and gaming in the English public health care system, Public Administration 84(3) (2006) 517-538.

[41] King's Fund, Have targets improved NHS performance? https://www.kingsfund.org.uk/projects/general-election-2010/performance-targets, 2010 (accessed April 21).

[42] G. Bevan, C. Hood, Have targets improved performance in the English NHS?, Bmj 332(7538) (2006) 419-422.

[43] M. Barber, Instruction to deliver: fighting to transform Britain's public services, Metheun, London, 2008.

[44] M.E. Barclay, G. Lyratzopoulos, D.C. Greenberg, G.A. Abel, Missing data and chance variation in public reporting of cancer stage at diagnosis: Cross-sectional analysis of population-based data in England, Cancer Epidemiol. 52 Suppl. C (2018) 28-42.

[45] Department of Health & Social Care, Handbook to the NHS Constitution for England, London, 2014.

[46] The King's Fund, An Alternative Guide to the Urgent and Emergency Care System in England, https://www.kingsfund.org.uk/audio-video/alternative-guide-urgent-and-emergency-care-system-england, 2015 (accessed 25 April 21).

[47] S. Benitez Majano, C. Di Girolamo, B. Rachet, *et al.,* Surgical treatment and survival from colorectal cancer in Denmark, England, Norway, and Sweden: a population-based study, Lancet Oncol. 20(1) (2019) 74-87.

[48] S. Walters, C. Maringe, M.P. Coleman, *et al.,* Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the United Kingdom: a population-based study, 2004-2007, Thorax 68 (2013) 551-564.

[49] J. Butler, C. Foot, M. Bomb, *et al.,* The International Cancer Benchmarking Partnership: an international collaboration to inform cancer policy in Australia, Canada, Denmark, Norway, Sweden and the United Kingdom, Health Policy 112(1-2) (2013) 148-55.

[50] M. Arnold, M.J. Rutherford, A. Bardot, *et al.,* Progress in cancer survival, mortality, and incidence in seven high-income countries 1995-2014 (ICBP SURVMARK-2): a population-based study, Lancet Oncol. 20(11) (2019) 1493-1505.

[51] Public Health England, Guidance: National Cancer Registration and Analysis Service (NCRAS), https://www.gov.uk/guidance/national-cancer-registration-and-analysis-service-ncras, 2019 (accessed 25 April 21).

[52] House of Commons Science and Technology Committee, Science and Technology - Sixth Report, London, 2000.

[53] National Disease Registration Service, FAQs: Does the cancer registry hold information about patients who are being treated privately?. https://www.ndrs.nhs.uk/question/does-the-cancer-registry-hold-information-about-patients-who-are-being-treated-privately/#:~:text=FAQs-,Does%20the%20cancer%20registry%20hold%20information%20about%20patients%20who%20are,data%20submitted%20to%20the%20NHS, (accessed 25 April 21).

[54] L. Elliss-Brookes, S. McPhail, A. Ives, *et al.,*, Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets, Br. J. Cancer 107(8) (2012) 1220-6.

[55] NHS England Contracts and Incentives Team, Quality Premium: Guidance for 2016/17, NHS England, Leeds, 2016.

[56] Office for National Statistics, Cancer Survival in England- Adults Diagnosed: 2009 to 2013, followed up to 2014, 2015.

[57] Public Health England, CancerData dashboard: Staging data in England (2020).

[58] Office for National Statistics, Cancer Registration Statistics, England: 2013, 2015.

[59] M. Astin, T. Griffin, R.D. Neal, P. Rose, W. Hamilton, The diagnostic value of symptoms for colorectal cancer in primary care: a systematic review, Br. J. Gen. Pract. 61(586) (2011) e231-e243.

[60] R.F.A. Logan, J. Patnick, C. Nickerson, L. Coleman, M.D. Rutter, C. von Wagner, Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests, Gut 61(10) (2011) 1439-46.

[61] J.K. Lee, E.G. Liles, S. Bent, T.R. Levin, D.A. Corley, Accuracy of fecal immunochemical tests for colorectal cancer: systematic review and meta-analysis, Ann. Intern. Med. 160(3) (2014) 171-171.

[62] S. Koo, L.J. Neilson, C. Von Wagner, C.J. Rees, The NHS Bowel Cancer Screening Program: current perspectives on strategies for improvement, Risk Manag. Healthc. Policy 10 (2017) 177-187.

[63] NHS England and NHS Improvement, Roll out of the new bowel cancer screening test - faecal immunochemical test (FIT) briefing for GPs, 2019.

[64] K.E. Nnoaham, A. Frater, P. Roderick, G. Moon, S. Halloran, Do geodemographic typologies explain variations in uptake in colorectal cancer screening? An assessment using routine screening data in the south of England, J. Public Health 32(4) (2010) 572-81.

[65] W.S. Atkin, R. Edwards, I. Kralj-Hans, *et al.,* Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial, Lancet 375(9726) 1624-1633.

[66] M.R. Thompson, K.G. Flashman, K. Wooldrage, *et al.,* Flexible sigmoidoscopy and whole colonic imaging in the diagnosis of cancer in patients with colorectal symptoms, Br. J. Surg. 95(9) (2008) 1140-1146.

[67] Information Centre for Health and Social Care, National Bowel Cancer Audit, Leeds, 2009.

[68] M.A. Richards, The size of the prize for earlier diagnosis of cancer in England, Br. J. Cancer 101 Suppl. 2 (2009) S125-9.

[69] Improved survival with preoperative radiotherapy in resectable rectal cancer. Swedish Rectal Cancer Trial, N. Engl. J. Med. 336(14) (1997) 980-7.

[70] E.J.A. Morris, P.J. Finan, K. Spencer, *et al.,* Wide Variation in the Use of Radiotherapy in the Management of Surgically Treated Rectal Cancer Across the English National Health Service, J. Clin. Oncol. 28(8) (2016) 522-531.

[71] J.E. van Hooft, E.E. van Halsema, G. Vanbiervliet, *et al.,* Self-expandable metal stents for obstructing colonic and extracolonic cancer: European Society of Gastrointestinal Endoscopy (ESGE) Clinical Guideline, Endosc. 46(11) (2014) 990-1053.

[72] R.S. Sawai, Management of colonic obstruction: a review, Clin. Colon. Rectal Surg. 25(4) (2012) 200-203.

[73] S.-H. Han, J.H. Lee, Colonic stent-related complications and their management, Clin. Endosc. 47(5) (2014) 415-419.

[74] E.J. Morris, E.F. Taylor, J.D. Thomas, *et al.,*Thirty-day postoperative mortality after colorectal cancer surgery in England, Gut 60(6) (2011) 806-813.

[75] S. Walters, S. Benitez Majano, P. Muller, *et al.,* Is England closing the international gap in cancer survival?, Br. J. Cancer. 113(5) (2015) 848-60.

[76] Cancer Research UK, Types of lung cancer. https://www.cancerresearchuk.org/about-cancer/lung-cancer/stages-types-grades/types, 2020 (accessed 25 April 21).

[77] Cancer Research UK, Mesothelioma incidence statistics. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/mesothelioma/incidence, 2020 (accessed 25 April 21).

[78] National Cancer Intelligence Network, Trends in incidence of small cell lung cancer and all lung cancers, London, 2010.

[79] National Health Service, Lung cancer symptoms. https://www.nhs.uk/conditions/lung-cancer/symptoms/, 2019 (accessed 25 April 21).

[80] E.L. O'Dowd, T.M. McKeever, D.R. Baldwin, *et al.,* What characteristics of primary care and patients are associated with early death in patients with lung cancer in the UK?, Thorax 70 (2015) 161-168.

[81] Cancer Research UK, Be Clear on Cancer evaluation update, National Health Service, London, 2014.

[82] J.K. Field, S.W. Duffy, D.R. Baldwin, *et al.,* The UK Lung Cancer Screening Trial: a pilot randomised controlled trial of low-dose computed tomography screening for the early detection of lung cancer, Health Technol. Assess. 20(40) (2016) 1-146.

[83] NHS England, The NHS Long Term Plan, 2019.

[84] Public Health England, Be Clear On Cancer – Overview. https://campaignresources.phe.gov.uk/resources/campaigns/16-be-clear-on-cancer/overview, 2019 (accessed 25 April 21).

[85] National Institute for Health and Clinical Excellence (NICE), Lung Cancer - The diagnosis and treatment of lung cancer, NICE clinical guideline, National Institute for Health and Clinical Excellence (NICE), 2011.

[86] M.D. Peake, Deprivation, distance and death in lung cancer, Thorax 70(2) (2015) 108-9.

[87] N. Patel, R. Adatia, A. Mellemgaard, R. Jack, H. Møller, Variation in the use of chemotherapy in lung cancer, Br. J. Cancer 96(6) (2007) 886-890.

[88] A. Khakwani, A.L. Rich, H.A. Powell, *et al.,* The impact of the 'hub and spoke' model of care for lung cancer and equitable access to surgery, Thorax 70(2) (2015) 146-51.

[89] H.A. Powell, L.J. Tata, D.R. Baldwin, R.A. Stanley, A. Khakwani, R.B. Hubbard, Early mortality after surgical resection for lung cancer: an analysis of the English National Lung cancer audit, Thorax 68(9) (2013) 826-834.

[90] M. Lüchtenborg, S.P. Riaz, V.H. Coupland, *et al.,* High procedure volume is strongly associated with improved survival after lung cancer surgery, J. Clin. Oncol. 31(25) (2013) 3141-3146.

[91] The Lancet, An experiment in earlier detection of ovarian cancer, Lancet 369(9579) (2007) 2051.

[92] B.A. Goff, L.S. Mandel, C.W. Drescher, *et al.,* Development of an ovarian cancer symptom index, Cancer 109(2) (2007) 221-227.

[93] C. Redman, S. Duffy, N. Bromham, K. Francis, Recognition and initial management of ovarian cancer: summary of NICE guidance, Bmj 342 (2011).

[94] National Institute for Health and Clinical Excellence (NICE), Ovarian cancer: recognition and initial management NICE guidelines, National Institute for Health and Clinical Excellence (NICE), 2011.

[95] S. Kehoe, J. Hook, M. Nankivell, *et al.,* Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): an open-label, randomised, controlled, non-inferiority trial, Lancet 386(9990) (2015) 249-257.

[96] M. Barclay, C. Gildea, J. Poole, L. Hirschowitz, U. Menon, A. Nordin, Factors Affecting Short-term Mortality in Women With Ovarian, Tubal, or Primary Peritoneal Cancer: Population-Based Cohort Analysis of English National Cancer Registration Data, Int. J. Gynecol. Cancer 26(1) (2016) 56-65.

[97] National Cancer Intelligence Network, Cancer survival in England by stage, 2014.

[98] D. Gupta, C.G. Lis, Pretreatment serum albumin as a predictor of cancer survival: A systematic review of the epidemiological literature, Nutr. J. 9(1) (2010) 69.

[99] J. Butler, C. Gildea, J. Poole, D. Meechan, A. Nordin, Specialist surgery for ovarian cancer in England, Gynecol. Oncol. 138(3) (2015) 700-6.

[100] R. Crawford, D. Greenberg, Improvements in survival of gynaecological cancer in the Anglia region of England: are these an effect of centralisation of care and use of multidisciplinary management?, BJOG 119(2) (2012) 160-165.

[101] Health & Social Care Information Centre, National Bowel Cancer Audit Data Specification v1.6, London, 2014.

[102] Royal College of Physicians, National Lung Cancer Audit. https://www.rcplondon.ac.uk/projects/national-lung-cancer-audit 2015 (accessed 25 April 21).

[103] M.E. Charlson, P. Pompei, K.L. Ales, C.R. MacKenzie, A new method of classifying prognostic comorbidity in longitudinal studies: development and validation, J. Chronic Dis. 40(5) (1987) 373-83.

[104] C. Maringe, H. Fowler, B. Rachet, M.A. Luque-Fernandez, Reproducibility, reliability and validity of population-based administrative health data for the assessment of cancer non-related comorbidities, PLoS One 12(3) (2017) e0172814.

[105] A. Boyd, R. Cornish, L. Johnson, *et al.,* Understanding Hospital Episode Statistics (HES), CLOSER, London, 2017.

[106] English indices of deprivation 2015 Ministry of Housing, Communities & Local Government, 2015.

[107] Office for National Statistics, Census Geography, 2016.

[108] J. Adams, M. White, Removing the health domain from the Index of Multiple Deprivation 2004—effect on measured inequalities in census measure of health, J. Public Health 28(4) (2006) 379-383.

[109] B. Rachet, C. Maringe, U. Nur, *et al.,* Population-based cancer survival trends in England and Wales up to 2007: an assessment of the NHS cancer plan for England, Lancet Oncol. 10(4) (2009) 351-69.

[110] A. Herbert, G.A. Abel, S. Winters, S. McPhail, L. Elliss-Brookes, G. Lyratzopoulos, Cancer diagnoses after emergency GP referral or A&E attendance in England: determinants and time trends in Routes to Diagnosis data, 2006–2015, Br. J. Gen. Pract. 69(687) (2019) e724-e730.

[111] S. Benitez Majano, H. Fowler, C. Maringe, C. Di Girolamo, B. Rachet, Deriving stage at diagnosis from multiple population-based sources: colorectal and lung cancer in England, Br. J. Cancer 115(3) (2016) 391-400.

[112] L. Sobin, M. Gospodarowicz, C. Wittekind, TNM Classification of Malignant Tumours, John Wiley & Sons, New York, 2009.

[113] S. McPhail, S. Johnson, D. Greenberg, M. Peake, B. Rous, Stage at diagnosis and early mortality from cancer in England, Br. J. Cancer 112(s1) (2015) S108-S115.

[114] Public Health England, Interpreting geographic variation in cancer stage: National Cancer Intelligence Data Briefing http://www.ncin.org.uk/publications/data_briefings/interpreting_geographic_variation_in_cancer_stage, 2014 (Accessed 25 April 21).

[115] C. Di Girolamo, S. Walters, S. Benitez Majano, *et al.,* Characteristics of patients with missing information on stage: a population-based study of patients diagnosed with colon, lung or breast cancer in England in 2013, BMC Cancer 18 (2018) 492.

[116] U.K.and Ireland Association of Cancer Registries, UKIACR Annual Performance Indicators: Commentaries, 2014.

[117] Cancer Research UK, United Kingdom and Ireland Association of Cancer Registries annual performance indicators for data published in 2015. https://www.ukiacr.org/sites/ukiacr/files/file-uploads/miscellaneous/UKIACR%20Performance%20Indicator%20Report%202015.pdf, 2015 (accessed 25 April 21).

[118] International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3), World Health Organization, 2000.

[119] R.J. Kurman, I.-M. Shih, Pathogenesis of ovarian cancer: lessons from morphology and molecular biology and their clinical implications, Int. J. Gynocol. Pathol. 27(2) (2008) 151-160.

[120] National Cancer Institute: SEER Training Modules, Morphology - Numerical List.

[121] National Cancer Institute: SEER Training Modules, Coding Rules for Topography and Morphology: Topography.

[122] National Health Service, Office of Population Censuses and Surveys Classification of Interventions and Procedures Version 4 (OPCS-4).

[123] Helen Fowler, Personal communication, Cancer Survival Group, London School of Hygiene & Tropical Medicine, 2019.

[124] M.A. Hernan, A definition of causal effect for epidemiological research, J. Epidemiol. Community Health 58(4) (2004) 265-71.

[125] Department of Health, Patient Reported Outcome Measures (PROMs) in England: The case-mix adjustment methodology, London, 2012.

[126] G.A. Abel, J. Shelton, S. Johnson, L. Elliss-Brookes, G. Lyratzopoulos, Cancer-specific variation in emergency presentation by sex, age and deprivation across 27 common and rarer cancers, Br. J. Cancer 112 Suppl. 1 (2015) 129-36.

[127] S. Walters, C. Maringe, J. Butler, *et al.*,, Breast cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK, 2000-2007: a population-based study, Br. J. Cancer 108 (2013) 1195-1208.

[128] B. Rachet, J. Siemiatycki, M. Abrahamowicz, K. Leffondre, A flexible modeling approach to estimating the component effects of smoking behavior on lung cancer, J. Clin. Epidemiol. 57(10) (2004) 1076-85.

[129] P. Murchie, E.A. Raja, D.H. Brewster, *et al.,* Time from first presentation in primary care to treatment of symptomatic colorectal cancer: effect on disease stage and survival, Br. J. Cancer 111(3) (2014) 461-9.

[130] R. Neal, Can earlier symptomatic diagnosis improve cancer outcomes in Wales? A report for Public Health Wales, Bangor, 2016.

[131] N.U. Din, O.C. Ukoumunne, G. Rubin, *et al.,* Age and Gender Variations in Cancer Diagnostic Intervals in 15 Cancers: Analysis of Data from the UK Clinical Practice Research Datalink, PLoS One 10(5) (2015) e0127717.

[132] A.E. Thompson, Y. Anisimowicz, B. Miedema, W. Hogg, W.P. Wodchis, K. Aubrey-Bassler, The influence of gender and other patient characteristics on health care-seeking behaviour: a QUALICOPC study, BMC Fam. Pract. 17 (2016) 38-38.

[133] C. Renzi, A. Kaushal, J. Emery, *et al.,* Comorbid chronic diseases and cancer diagnosis: disease-specific effects and underlying mechanisms, Nat. Rev. Clin. Oncol. 16(12) (2019) 746-761.

[134] S. Quadrelli, G. Lyons, H. Colt, D. Chimondeguy, A. Buero, Clinical Characteristics and Prognosis of Incidentally Detected Lung Cancers, Int. J. Surg. Oncol. (2015) 287604.

[135] C.A. Welch, M.J. Sweeting, P.C. Lambert, *et al.,* Impact on survival of modelling increased surgical resection rates in patients with non-small-cell lung cancer and cardiovascular comorbidities: a VICORI study, Br. J. Cancer 123(3) (2020) 471-479.

[136] G. Lyratzopoulos, J.M. Barbiere, B. Rachet, M. Baum, M.R. Thompson, M.P. Coleman, Changes over time in socioeconomic inequalities in breast and rectal cancer survival in England and Wales during a 32-year period (1973-2004): the potential role of health care, Ann. Oncol. 22(7) (2011) 1661-6.

[137] D.H. Brewster, C.S. Thomson, D.J. Hole, R.J. Black, P.L. Stroner, C.R. Gillis, Relation between socioeconomic status and tumour stage in patients with breast, colorectal, ovarian, and lung cancer: results from four national, population based studies, Bmj 322(7290) (2001) 830-1.

[138] M.E. Barclay, G.A. Abel, D.C. Greenberg, B. Rous, G. Lyratzopoulos, Socio-demographic variation in stage at diagnosis of breast, bladder, colon, endometrial, lung, melanoma, prostate, rectal, renal and ovarian cancer in England and its population impact, British Journal of Cancer 124(7) (2021) 1320-1329.

[139] G. Lyratzopoulos, G.A. Abel, C.H. Brown, B.A. Rous, S.A. Vernon, M. Roland, D.C. Greenberg, Socio-demographic inequalities in stage of cancer diagnosis: evidence from patients with female breast, lung, colon, rectal, prostate, renal, bladder, melanoma, ovarian and endometrial cancer, Ann Oncol 24(3) (2013) 843-50.

[140] R. Raine, W. Wong, S. Scholes, C. Ashton, A. Obichere, G. Ambler, Social variations in access to hospital care for patients with colorectal, breast, and lung cancer between 1999 and 2006: retrospective analysis of hospital episode statistics, Bmj 340 (2010) b5479.

[141] House of Commons Committee of Public Accounts, Tackling inequalities in life expectancy in areas with the worst health and deprivation, Third report of Session 2010-11, London, 2010.

[142] R.A. Hahn, B.I. Truman, Education improves public health and promotes health equity, Int. J. Health Serv. 45(4) (2015) 657-678.

[143] P. Savage, R. Sharkey, T. Kua, *et al.,* Clinical characteristics and outcomes for patients with an initial emergency presentation of malignancy: a 15 month audit of patient level data, Cancer Epidemiol. 39(1) (2015) 86-90.

[144] The Global Health Observatory, Age-standardized mortality rate (per 100 000 population), World Health Organisation, 2020.

[145] C. Allemani, T. Matsuda, V. Di Carlo, *et al.,* Global surveillance of trends in cancer survival 2000-2014 (CONCORD-3): analysis of individual records for 37,513,025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries, Lancet 391(10125) (2018) 1023-1075.

[146] J. Fletcher, Standardised mortality ratios, Bmj 338 (2009) b2005.

[147] D.J. Spiegelhalter, Funnel plots for comparing institutional performance, Stat. Med. 24 (2005) 1185-1202.

[148] B. Kirkwood, J. Sterne, Essential Medical Statistics, second ed., Wiley, Hoboken, 2003.

[149] M. Inkster, A. Montgomery, P. Donnan, T. MacDonald, F. Sullivan, T. Fahey, Organisational factors in relation to control of blood pressure: an observational study, Br. J. Gen. Pract. 55(521) (2005) 931-937.

[150] S. Greenland, J.M. Robins, J. Pearl, Confounding and collapsibility in causal inference, Stat. Sci. 14(1) (1999) 29-46.

[151] M.A. Hernán, D. Clayton, N. Keiding, The Simpson's paradox unraveled, Int. J. Epidemiol. 40(3) (2011) 780-785.

[152] A.D. Sinaiko, A.T. Chien, M.J. Hassett, *et al.,* What drives variation in spending for breast cancer patients within geographic regions?, Health Serv. Res. 54(1) (2019) 97-105.

[153] P.C. Austin, J. Merlo, Intermediate and advanced topics in multilevel logistic regression analysis, Stat. Med. 36(20) (2017) 3257-3277.

[154] J.T. Lee, Z. Huang, S. Basu, C. Millett, The inverse equity hypothesis: Does it apply to coverage of cancer screening in middle-income countries?, J. Epidemiol. Community Health 69(2) (2015) 149.

[155] C.G. Victora, G. Joseph, I.C.M. Silva, *et al.,* The Inverse Equity Hypothesis: Analyses of institutional deliveries in 286 national surveys, Am. J. Public Health 108(4) (2018) 464-471.

[156] M. Quaresma, J. Jenkins, N. Bannister, *et al.,* Index of cancer survival for Clinical Commissioning Groups in England: adults diagnosed 1999 to 2014 and followed up to 2015, Office for National Statistics, 2016.

[157] Department of Health, Healthy Lives, Healthy People: Improving Outcomes and Supporting Transparency. Part 1: A Public Health Outcomes Framework for England, 2013-2016, Stationery Office, London, 2013, 22.

[158] P. Muller, L. Woods, S. Walters, Temporal and geographic changes in stage at diagnosis in England during 2008–2013: A population-based study of colorectal, lung and ovarian cancers, Cancer Epidemiol. 67 (2020) 101743.

[159] D. Spiegelhalter, C. Sherlaw-Johnson, M. Bardsley, I. Blunt, C. Wood, O. Grigg, Statistical methods for healthcare regulation: rating, screening and surveillance, J. R. Stat. Soc. Ser. A. Stat. Soc. 175(1) (2012) 1-47.

[160] Healthcare Quality Improvement Partnership, National Bowel Cancer Audit report 2014, Leeds, 2014.

[161] D.J. Spiegelhalter, Personal communication, 2016.

[162] S. Rabe-Hesketh, A. Skrondal, Multilevel and Longitudinal Modeling Using Stata., third ed., Stata Press, College Station, 2012.

[163] G. Lyratzopoulos, M.N. Elliott, J.M. Barbiere, *et al.,* How can health care organizations be reliably compared?: Lessons from a national survey of patient experience, Med. Care 49(8) (2011) 724-33.

[164] J.L. Schafer, J.W. Graham, Missing data: our view of the state of the art, Psychol. Methods 7(2) (2002) 147-77.

[165] D.B. Rubin, Multiple Imputation For Nonresponse In Surveys, John Wiley and Sons, New York, 1987.

[166] I.R. White, P. Royston, A.M. Wood, Multiple imputation using chained equations: Issues and guidance for practice, Stat. Med. 30(4) (2011) 377-399.

[167] M. Falcaro, U. Nur, B. Rachet, J.R. Carpenter, Estimating excess hazard ratios and net survival when covariate data are missing: strategies for multiple imputation, Epidemiology 26(3) (2015) 421-8.

[168] J.C. Jakobsen, C. Gluud, J. Wetterslev, P. Winkel, When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts, BMC Medical Research Methodology 17(1) (2017) 162.

[169] P. Muller, S. Walters, M.P. Coleman, L. Woods, Which indicators of early cancer diagnosis from population-based data sources are associated with short-term mortality and survival?, Cancer Epidemiol. 56 (2018) 161-170.

[170] T.E. Bodner, What improves with increased missing data imputations?, Struct. Equ. Modeling 15(4) (2008) 651-675.

[171] J.H. Lee, J. Huber, Multiple imputation with large proportions of missing data: How much is too much?, United Kingdom Stata User's Group Meeting, 2011.

[172] M. Quartagno, J. Carpenter, jomo: A package for Multilevel Joint Modelling Multiple Imputation, 2020 https://cran.r-project.org/package=jomo.

[173] R.J. Kurman, I.-M. Shih, The dualistic model of ovarian carcinogenesis: Revisited, revised, and expanded, Am. J. Pathol. 186(4) (2016) 733-747.

[174] M. Matz, M.P. Coleman, M. Sant, *et al.,* The histology of ovarian cancer: worldwide distribution and implications for international survival comparisons (CONCORD-2), Gynecol. Oncol. 144(2) (2017) 405-413.

[175] J.W. Graham, A.E. Olchowski, T.D. Gilreath, How many imputations are really needed? Some practical clarifications of multiple imputation theory, Prev. Sci. 8(3) (2007) 206-13.

[176] StataCorp, meqrlogit - Multilevel mixed-effects logistic regression (QR decomposition), 2013.

[177] M.B. Larsen, S. Njor, P. Ingeholm, B. Andersen, Effectiveness of colorectal cancer screening in detecting earlier-stage disease—A nationwide cohort study in Denmark, Gastroenterology 155(1) (2018) 99-106.

[178] J.-B. Lew, D.J.B. St John, X.-M. Xu, *et al.,* Long-term evaluation of benefits, harms, and cost-effectiveness of the National Bowel Cancer Screening Program in Australia: a modelling study, Lancet Public Health 2(7) (2017) e331-e340.

[179] Public Health England, National Cancer Registration and Analysis Service, Routes to Diagnosis 2006-2016 year breakdown. http://www.ncin.org.uk/publications/routes_to_diagnosis, 2018 (accessed 25 April 21).

[180] H. Brenner, B. Rachet, Hybrid analysis for up-to-date long-term survival rates in cancer registries with delayed recording of incident cases, Eur. J. Cancer 40(16) (2004) 2494-501.

[181] Public Health England, COVID-19: Diagnosis and treatment information for cancer patients in England during the COVID-19 pandemic. https://www.cancerdata.nhs.uk/covid-19, 2021 (accessed 25 April 21).

[182] L. Ironmonger, E. Ohuma, N. Ormiston-Smith, C. Gildea, C.S. Thomson, M.D. Peake, An evaluation of the impact of large-scale interventions to raise public awareness of a lung cancer symptom, Br J Cancer 112(1) (2015) 207-16.

[183] M. Falcaro, J.R. Carpenter, Correcting bias due to missing stage data in the non-parametric estimation of stage-specific net survival for colorectal cancer using multiple imputation, Cancer Epidemiol. 48 (2017) 16-21.

[184] N. Eisemann, A. Waldmann, A. Katalinic, Imputation of missing values of tumour stage in population-based cancer registration, BMC Med. Res. Methodol. 11(1) (2011) 129.

[185] Q. Luo, S. Egger, X.Q. Yu, D.P. Smith, D.L. O'Connell, Validity of using multiple imputation for "unknown" stage at diagnosis in population-based cancer registry data, PLoS One 12(6) (2017) e0180033.

[186] C. Maringe, J. Spicer, M. Morris, *et al.,* The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study, Lancet Oncol. 21(8) (2020) 1023-1034.

[187] T.M. Andersson, M.J. Rutherford, K. Humphreys, Assessment of lead-time bias in estimates of relative survival for breast cancer, Cancer Epidemiol 46 (2017) 50-56.

[188] K. Hemming, T.P. Haines, P.J. Chilton, A.J. Girling, R.J. Lilford, The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting, BMJ : British Medical Journal 350 (2015) h391.