International Journal of Population Data Science

Journal Website: www.ijpds.org

Variation in colorectal cancer treatment and outcomes in Scotland: real world evidence from national linked administrative health data

Elizabeth Lemmon[1,5,*], Catherine Hanna[2], Katharina Diernberger[1,5], Hugh M. Paterson[3], Sarah H. Wild[6], Holly Ennis[5], and Peter S. Hall[1,4]

Submission History

Submitted:	06/07/2023
Accepted:	20/12/2023
Published:	20/02/2024

[1]Edinburgh Health Economics, University of Edinburgh
[2]School of Cancer Sciences, University of Glasgow
[3]Department of Colorectal Surgery, Western General Hospital, NHS Lothian, Edinburgh; University of Edinburgh
[4]Edinburgh Cancer Research Centre, University of Edinburgh
[5]Edinburgh Clinical Trials Unit, University of Edinburgh
[6]Usher Institute, University of Edinburgh

Abstract

Background

Colorectal cancer (CRC) is the fourth most common type of cancer in the United Kingdom and the second leading cause of cancer death. Despite improvements in CRC survival over time, Scotland lags behind its UK and European counterparts. In this study, we carry out an exploratory analysis which aims to provide contemporary, population level evidence on CRC treatment and survival in Scotland.

Methods

We conducted a retrospective population-based analysis of adults with incident CRC registered on the Scottish Cancer Registry (Scottish Morbidity Record 06 (SMR06)) between January 2006 and December 2018. The CRC cohort was linked to hospital inpatient (SMR01) and National Records of Scotland (NRS) deaths records allowing a description of their demographic, diagnostic and treatment characteristics. Cox proportional hazards regression models were used to explore the demographic and clinical factors associated with all-cause mortality and CRC specific mortality after adjusting for patient and tumour characteristics among people identified as early-stage and treated with surgery.

Results

Overall, 32,691 (73%) and 12,184 (27%) patients had a diagnosis of colon and rectal cancer respectively, of whom 55% and 53% were early-stage and treated with surgery. Five year overall survival (CRC specific survival) within this cohort was 72% (82%) and 76% (84%) for patients with colon and rectal cancer respectively. Cox proportional hazards models revealed significant variation in mortality by sex, area-based deprivation and geographic location.

Conclusions

In a Scottish population of patients with early-stage CRC treated with surgery, there was significant variation in risk of death, even after accounting for clinical factors and patient characteristics.

Keywords

colorectal cancer; survivorship; administrative data; geographic variation; cancer registry

*Corresponding Author:
Email Address: elizabeth.lemmon@ed.ac.uk (Elizabeth Lemmon)

https://doi.org/10.23889/ijpds.v9i1.2179

Introduction

Colorectal cancer (CRC) is the fourth most common type of cancer for men and women in the United Kingdom (UK) and the second leading cause of cancer death [1]. Compared to the rest of the UK, bowel cancer incidence is highest in Scotland [2] and it is projected that the number of new CRC cases in Scotland will increase by 43% by 2023-27 compared to 2008-12 [3]. Since the introduction of the Scottish Bowel Screening Programme in June 2007, uptake of screening has continued to increase to an all-time high of 65% during the period 1st May 2019 to 30th April 2021 [4]. The programme aims to detect CRC earlier and even prevent CRC occurring in the first place by identifying and removing pre-cancerous adenomatous polyps.

Despite improvements in CRC survival over time, Scotland lags behind its UK and European counterparts [5, 6]. Previous research has shown that socio-economic deprivation and remoteness factors, for example distance from a cancer centre, are significantly associated with poorer survival in Scotland [7, 8] and in the UK [9–11]. However, there are considerable differences in conclusions across studies and the mechanisms behind the observed relationships remain unclear [12–14].

Research in this area, and ultimately patient outcomes, may be improved by utilising the vast amounts of administrative healthcare data that are collected routinely as part of the delivery of patient care [16]. These data provide an opportunity to generate evidence with a high degree of external validity, being entirely representative of current care [17].

In England, linked administrative datasets have been used to investigate routes to CRC diagnosis [18]; explore provider differences in post-colonoscopy CRC diagnosis rates [19]; explain variation in treatment and outcomes [20, 21]; and describe management of disease [22]. In Scotland, there are very few CRC studies using linked administrative datasets to investigate CRC treatment and many have used administrative data from a single geographic area e.g. [12, 14, 23]. Furthermore, published, population level statistics on colorectal cancer survival [15] group stages and/or disease sites together, without distinguishing patients who are treated at early-stage or accounting for important prognostic factors.

In this study, we contribute to the existing research by providing a contemporary, population level description of the demographic, diagnostic and treatment characteristics of patients diagnosed with CRC in Scotland. We also carry out an exploratory analysis of the factors associated with survival for those early-stage, surgically treated patients. We use a newly established, unique CRC dataset, which links demographic data to the Scottish Cancer Registry and routinely collected hospital admissions data. Full details of the linkage and creation of this source dataset are described elsewhere [24]. Public Health Scotland's (PHS) Data Quality Assurance team carry out regular data accuracy audits on national datasets. Unlike other published work, we focus on patients defined as early-stage disease treated with surgery and incorporate prognostic factors such as disease stage and comorbidities. In what follows, we firstly describe the demographic, diagnostic and treatment characteristics of patients with colon and rectal cancer. Secondly, for those patients identified as early-stage, surgically treated, we estimate their survival and explore the factors affecting their overall survival and CRC specific survival.

Methods

The study design is a retrospective cohort study of adults with an incident CRC registered on the Scottish Cancer Registry (International Disease Classification 10th Revision (ICD-10) codes C18, C19 and 20) between January 2006 and December 2018. Approval for the study was granted by the Public Benefit and Privacy Panel (PBPP) for health and social care, project number 1718-0026. The study meets the requirements set out by the East of Scotland NHS Research Ethics Service for the analysis of secondary National Services of Scotland (NSS) data [25]. This study follows the REporting of studies Conducted using Observational Routinely collected health Data (RECORD) checklist [26].

Data

We used the Scottish Cancer Registry (Scottish Morbidity Record 06 (SMR06)) to identify a cohort of patients diagnosed with CRC. Linked National Records of Scotland (NRS) data was used to provide date and cause of death and inpatient and day case hospital admissions data (SMR01) provided information on comorbidities and hospital use prior to cancer diagnosis.

These three datasets were linked via a pseudonomysied patient identifier.

Scottish cancer registry (SMR06)

The Scottish Cancer Registry dataset includes information on all diagnoses of cancer occurring within Scotland. The data is collected by PHS and contains diagnostic, staging and treatment information. Each SMR06 record for a patient corresponds to a unique cancer diagnosis for that individual.

In this study, we had access to all SMR06 records for patients who had a diagnosis of CRC (ICD-10 codes C18, C19 or C20) between January 2006 and December 2018 including multiple SMR06 records for people with more than one CRC diagnosis between 1981 and 2006 and people with a non-CRC diagnosis during the study period.

Inpatient and day case admissions (SMR01)

The SMR01 dataset contains episode level data for all general/acute inpatient or day cases in Scottish NHS hospitals or Scottish NHS beds in non-NHS-institutions. This study used all patient SMR01 records for the five years preceding diagnosis of CRC, for any patient present in the study SMR06 dataset.

National records of Scotland (Deaths)

The NRS is responsible for the registration of all life events occurring in Scotland including births, deaths, marriages, civil partnerships and adoptions. For the purposes of this study, NRS vital events data on births and deaths were used to obtain patient date of birth, sex, date and cause of death prior to the end of the study period.

Cohort derivation

The retrospective cohort was derived from the SMR06 database. Prior to linking SMR06 records to SMR01 and deaths records, a number of exclusion criteria were applied. These criteria are described in Figure 1.

Following exclusions, a total of 44,875 records corresponding to unique individuals remained in the cohort. This cohort was then linked to the SMR01 records to obtain five-year pre-CRC diagnosis comorbidity and hospital admissions information, and finally to deaths records to obtain survival outcomes.

Descriptive analysis

The full cohort was characterised by descriptive statistics of their demographics, diagnosis and treatment. The variables included are described along with their sources in Table 1 below.

Since disease trajectories and treatment pathways are different for patients with rectal cancer (ICD.10 code C20) from patients with colon cancer (ICD-10 codes C18 and C19), analysis was carried out separately for these two disease sites.

Statistical analysis

For the statistical analysis, we focus on patients treated with surgery and at an early-stage due to the very different treatment pathways for those diagnosed with late-stage disease. After removing those who did not undergo surgery (or who were coded as unknown or planned surgery) (n = 11,689); those who were diagnosed with tumour stage IV or unknown disease stage (n = 4969); those who had palliative or unknown therapy objectives (n = 3,679) (who are most likely to be more severe stage three individuals or those with complications/other comorbidities we cannot capture); and those with missing chemotherapy and radiotherapy treatment variables (n = 156), a final early-stage surgically treated cohort of 24,383 patients remained.

We defined two end points: Overall survival (OS) and CRC survival (CRCS). OS was the interval between the date of diagnosis of CRC and the date of death, censored at 31st December 2018 for survivors to this point. CRCS was defined as the interval between the date of diagnosis and the date of death, from CRC as the underlying cause of death. Survivors were censored in the same way as outlined for OS with the additional censoring at date of death for those who died from non-CRC related causes. Survival curves were estimated using the Kaplan-Meier method. Survival outcomes are reported for three-, five- and ten-year time points.

Figure 1: Cohort derivation flow chart

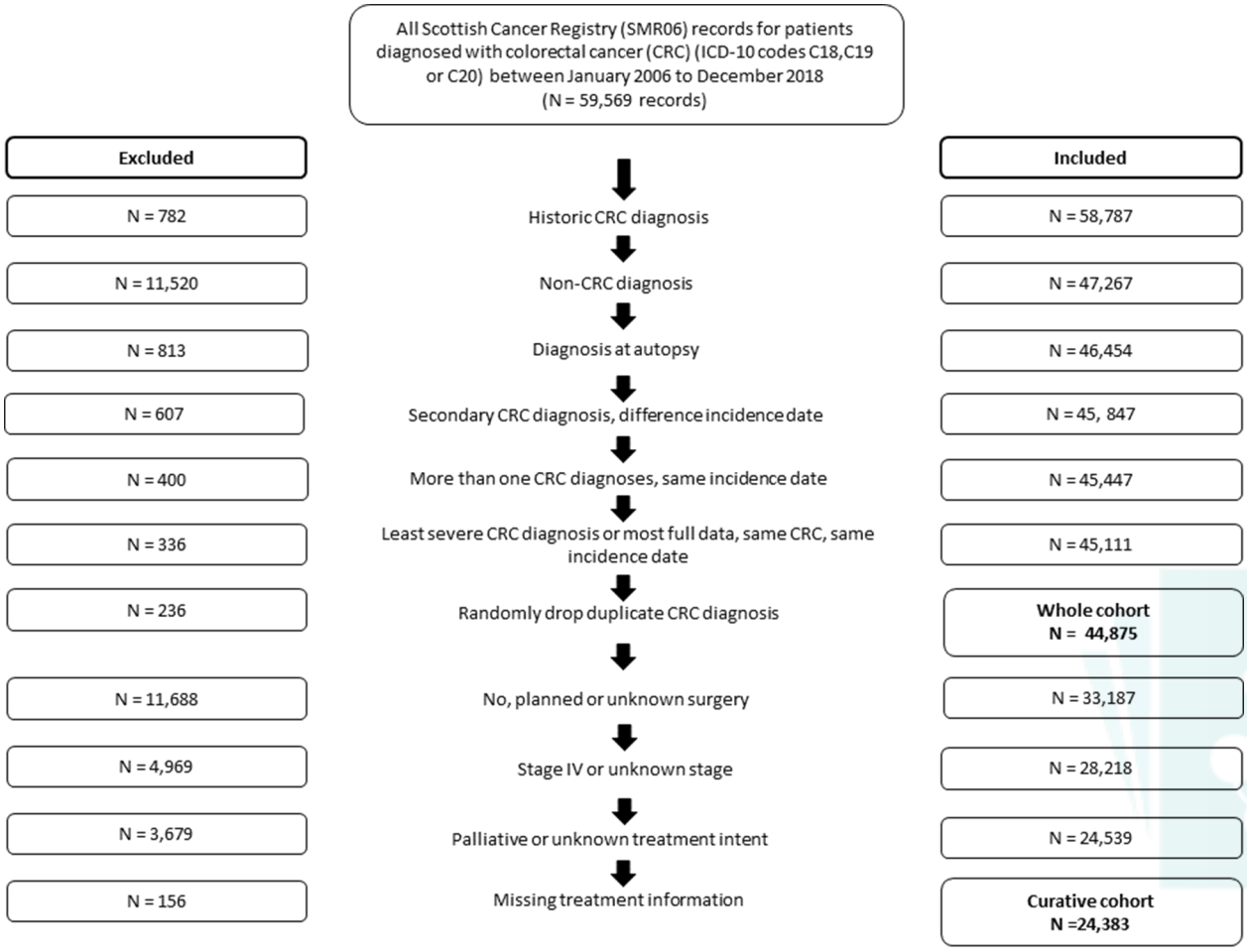


Table 1: Definitions of variables

Variable	Description	Data source
Demographics		
Sex	Binary sex indicator to indicate if the patient is male or female	SMR06
Age	10-year age bands (18-34, 35-44, 45-54, 55-64, 65-74, 75-84, 85+) based on age at diagnosis	SMR06
SIMD Quintile	Scottish Index of Multiple Deprivation (SIMD) quintile from most deprived (1) to least deprived (5)	Scottish Government
Urban/Rural	Binary indicator to indicate the rurality of the usual residence of the patient. Comes from the Scottish Government's six-fold urban/rural classification. Urban locations are defined as large urban areas, other urban areas or accessible small towns. Remote/rural are defined as remote small towns, accessible rural or remote rural.	Scottish Government
Cancer Network	An indicator of the cancer Managed Clinical Network (MCN) in which the patient was diagnosed. There are three MCNs in Scotland including South of Scotland Cancer Network (SCAN), West of Scotland Cancer Network (WoSCAN) and the North of Scotland Cancer Network (NoSCAN)	SMR06
Diagnosis		
Tumour Stage	Categorical indicator corresponding to Duke's stage indicating the extent of spread of the invasive tumour at diagnosis in terms of the pathological and/or clinical findings. Where Duke's stage is missing, pathologic American Joint Committee on Cancer (AJCC) stage is used. Where pathologic stage is missing, clinical AJCC stage is used. (Stage I,II,III, IV or Unknown)	SMR06
Method 1st detection	Categorical indicator to indicate how the tumour was first detected (Screening examination, incidental finding, clinical presentation, interval cancer or other, not known).	SMR06
Year of diagnosis	Three year bands for the year of diagnosis (2006-08, 2009-11, 2012-14 and 2015-18)	SMR06
Treatment		
Therapy objectives	Categorical indicator to indicate the treatment intent (curative, palliative or unknown)	SMR06
Chemotherapy	Binary indicator to indicate if the patient has had systemic chemotherapy treatment (1 if yes, 0 if no)	SMR06
Radiotherapy	Binary indicator to indicate if the patient was treated with radiotherapy (1 if yes, 0 if no)	SMR06
Surgery	Binary indicator to indicate if the patient was treated with surgery (1 if yes, 0 if no)	SMR06
Inpatient episodes	Mean number of hospital inpatient episodes or day cases in the five years pre-diagnosis	SMR01
Comorbidity score	Mean comorbidity score [27] in the five years pre-diagnosis using Charlson indicators with Quan weights (excludes cancer and metastatic cancer).	SMR01

Finally, we conducted univariable and multivariable analysis to examine the factors associated with all-cause mortality and CRC specific mortality using Cox proportional hazards models. The assumption of proportional hazards was checked via a visual inspection of the Kaplan-Meier curves and statistical tests of the Schoenfield residuals. In order to account for differences in patient characteristics and to capture differences across groups and location, multivariable models were adjusted for sex, age group, cancer network, Scottish Index of Multiple Deprivation (SIMD) quintile, tumour stage, urban/rural, year of diagnosis, number of inpatient episodes and comorbidity score in the five years pre-diagnosis [27].

Results

The final study population included 44,875 patients and the early-stage surgically treated cohort included 24,383 patients. Patient characteristics are described in Table 2. Overall, 32,691 (73%) of patients had a diagnosis of colon cancer and 12,184 (27%) had a diagnosis of rectal cancer. Of those patients diagnosed with colon cancer, 55% were included in the early-stage surgically treated cohort compared to 53% of those diagnosed with rectal cancer. Supplementary Table A, Table B and Table C display the characteristics of patients by cancer network, prior to the early-stage surgically treated cohort selection.

Table 3 presents survival outcomes for the early-stage surgically treated cohort (n = 24,383) and Figure 2 presents the Kaplan-Meier (KM) survival curves for OS and CRC.

Tables 4 and 5 display the results of Cox proportional hazard models. Adjusted models for all-cause mortality and CRC specific mortality show that females have a significantly better survival compared to males. Survival was inversely associated with deprivation. For both cancers, compared to patients in the most deprived quintile, those in the least deprived quintile have a significantly reduced risk of death. Survival improved over time, with diagnoses in later years associated with better outcomes. Further, increased risk of death was observed as comorbidity increases. The results also show that patients diagnosed with colon cancer in SCAN have significantly increased risk of death from all causes and from CRC relative to patients in WoSCAN.

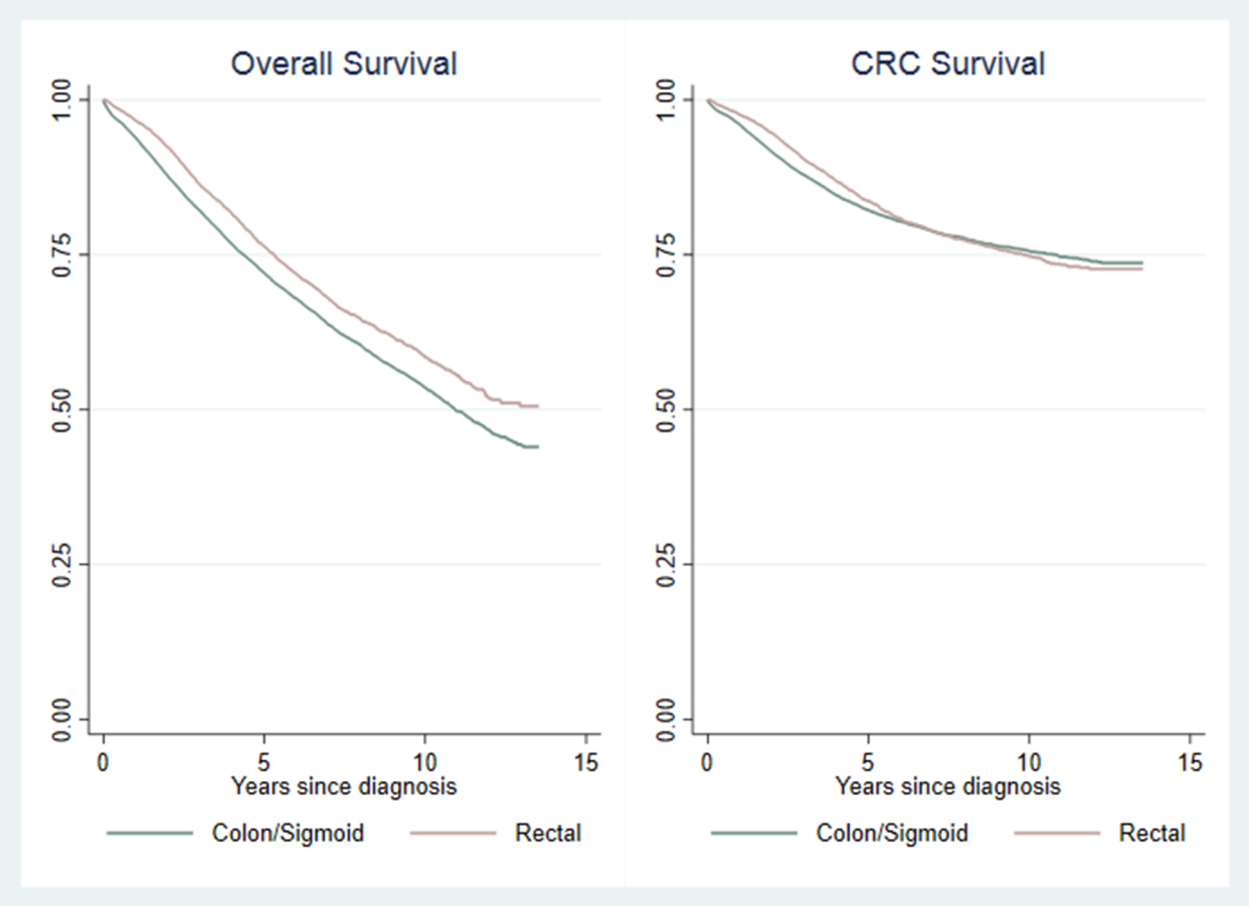
Table 2: Characteristics of patients diagnosed with colon or rectal cancer in Scotland between January 2006 to December 2018

		Total cohort N = 44,875				Curative cohort N = 24,383			
		Colon/Sigmoid		Rectal		Colon/Sigmoid		Rectal	
		N = 32,691		N = 12,184		N = 17,943		N = 6,440	
		n	%	n	%	n	%	n	%
Demographics									
Sex	Male	16,989	52.0	7,532	61.8	9,447	52.7	4,030	62.6
	Female	15,702	48.0	4,652	38.2	8,496	47.3	2,410	37.4
Age	18–34	265	0.8	91	0.7	154	0.9	45	0.7
	35–44	598	1.8	325	2.7	325	1.8	193	3.0
	45–54	2,360	7.2	1,227	10.1	1,402	7.8	751	11.7
	55–64	5,722	17.5	2,754	22.6	3,462	19.3	1,647	25.6
	65–74	9,923	30.4	3,776	31.0	6,140	34.2	2,256	35.0
	75–84	9,947	30.4	2,997	24.6	5,210	29.0	1,327	20.6
	85+	3,876	11.9	1,014	8.3	1,250	7.0	221	3.4
SIMD Quintile	1	6,296	19.3	2,368	19.4	3,237	18.0	1,123	17.4
	2	6,994	21.4	2,634	21.6	3,785	21.1	1,337	20.8
	3	7,002	21.4	2,516	20.7	3,881	21.6	1,319	20.5
	4	6,247	19.1	2,436	20.0	3,466	19.3	1,390	21.6
	5	6,152	18.8	2,230	18.3	3,574	19.9	1,271	19.7
Urban/Rural	Urban	25,018	76.5	9,242	75.9	13,748	76.6	4,867	75.6
	Remote or Rural	7,673	23.5	2,942	24.1	4,195	23.4	1,573	24.4
Cancer Network	SCAN	11,344	34.7	4,110	33.7	6,871	38.3	2,350	36.5
	WoSCAN	13,186	40.3	4,880	40.1	6,823	38.0	2,441	37.9
	NoSCAN	8,161	25.0	3,194	26.2	4,249	23.7	1,649	25.6
Diagnosis									
Tumour stage	I	4,139	12.7	2,651	21.8	3,982	22.2	2,510	39.0
	II	8,909	27.3	2,310	19.0	8,315	46.3	2,102	32.6
	III	8,035	24.6	2,524	20.7	5,646	31.5	1,828	28.4
	IV	7,169	21.9	1,947	16.0	–	–	–	–
	Unknown	4,439	13.6	2,752	22.6	–	–	–	–
Method 1st detection	Screening	4,432	13.6	1,799	14.8	3,641	20.3	1,399	21.7
	Incidental finding	983	3.0	206	1.7	448	2.5	83	1.3
	Clinical presentation	26,912	82.3	10,065	82.6	13,737	76.6	4,925	76.5
	Interval cancer, other	152	0.5	32	0.3	152	0.5	32	0.3
	Unknown	212	0.6	82	0.7	44	0.6	82	0.7
Year of diagnosis	2006–2008	7,552	23.1	2,941	24.1	3,719	20.7	1,338	20.8
	2009–2011	8,298	25.4	3,094	25.4	4,688	26.1	1,692	26.3
	2012–2014	7,986	24.4	2,823	23.2	4,546	25.3	1,586	24.6
	2015–2018	8,855	27.1	3,326	27.3	4,990	27.8	1,824	28.3
Treatment									
Therapy objectives	Curative	19,312	59.1	7,585	62.3	17,943	100.0	6,440	100.0
	Palliative	11,403	34.9	3,659	30.0	–	–	–	–
	Unknown	1,976	6.0	940	7.7	–	–	–	–
Chemotherapy	No	22,723	69.5	6,922	56.8	12,929	72.1	3,845	59.7
	Yes/planned	9,654	29.5	5,136	42.2	5,014	27.9	2,595	40.3
	Not known	314	1.0	126	1.0	–	–	–	–
Radiotherapy	No	31,556	96.5	7,205	59.1	17,747	98.9	4,300	66.8
	Yes/planned	767	2.3	4,794	39.3	196	1.1	2,140	33.2
	Not known	368	1.1	185	1.5	–	–	–	–
Surgery	No	8,145	24.9	3,360	27.6	–	–	–	–
	Yes/planned	24,436	74.7	8,765	71.9	17,943	100.0	6,440	100.0
	Not known	110	0.3	59	0.5	–	–	–	–
Inpatient episodes	Mean	1.99	–	1.95	–	1.98	–	1.96	–
Comorbidity score	Mean	0.07	–	0.06	–	0.05	–	0.05	–

Table 3: Survival Outcomes for early-stage patients diagnosed with colon or rectal cancer in Scotland between January 2006 and December 2018 and treated with surgery

	Early-stage surgically treated cohort N = 24,383					
	EAll deaths			CRC deaths		
	All	Colon/Sigmoid	Rectal	All	Colon/Sigmoid	Rectal
Number of patients	24,383	17,943	6,440	24,383	17,943	6,440
Number of deaths	8,102	6,184	1,918	4,348	3,213	1,135
Median survival (years)	11.42	10.93	–	–	–	–
3-year survival	83%	82%	86%	88%	88%	90%
5-year survival	73%	72%	76%	83%	82%	84%
10-year survival	55%	54%	59%	75%	76%	75%

Figure 2: Kaplan-Meier Survival Curves by CRC type for patients diagnosed with early-stage colon or rectal cancer in Scotland between January 2006 and December 2018 and treated with surgery



Discussion

This paper used a national linked administrative dataset to describe factors associated with survival for patients diagnosed with CRC. We described the demographic, diagnostic and treatment characteristics of patients with colon cancer and rectal cancer. Further, for early-stage patients treated with surgery, we assessed the factors affecting overall and CRC specific survival.

The survival models concur with existing UK and other worldwide evidence in a number of areas. Specifically, Cox-proportional hazard models for an early-stage surgically treated cohort of patients with colon or rectal cancer confirm that, stage at diagnosis strongly influences mortality/survival, with those diagnosed at more advanced disease stage having poorer survival compared to those diagnosed at the earliest stage. The adjusted models for deaths from all causes show that patients diagnosed with stage III colon or rectal cancer have a twofold increase in risk of death when compared to those diagnosed at stage I. For CRC specific deaths, this risk increases to almost

Table 4: Cox-proportional hazard regression models: patients diagnosed with early-stage colon cancer in Scotland between January 2006 and December 2018 and treated with surgery, n = 17,941

Variable	Category	All-cause mortality						CRC-specific mortality					
		Unadjusted			Adjusted			Unadjusted			Adjusted		
		HR	95 %	CI	HR	95 %	CI	HR	95 %	CI	HR	95%	CI
Sex	Male	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	Female	**0.92**	**0.87**	**0.97**	**0.84**	**0.80**	**0.88**	0.93	0.87	1.00	**0.87**	**0.82**	**0.94**
Age (years)	18–34	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	35–44	1.22	0.69	2.14	1.05	0.59	1.85	1.38	0.78	2.43	1.11	0.60	2.04
	45–54	1.19	0.72	1.98	1.09	0.66	1.81	1.26	0.76	2.08	1.09	0.63	1.89
	55–64	1.56	0.95	2.55	1.47	0.90	2.41	1.39	0.85	2.27	1.28	0.75	2.20
	65–74	**2.30**	**1.41**	**3.76**	**2.23**	**1.36**	**3.64**	1.57	0.96	2.56	1.52	0.89	2.60
	75–84	**4.71**	**2.89**	**7.69**	**4.57**	**2.80**	**7.46**	**2.67**	**1.64**	**4.36**	**2.60**	**1.53**	**4.43**
	85+	**8.31**	**5.08**	**13.59**	**8.15**	**4.98**	**13.35**	**3.99**	**2.44**	**6.53**	**3.89**	**2.27**	**6.68**
Cancer Network	WoSCAN	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	SCAN	**1.11**	**1.05**	**1.18**	**1.06**	**1.00**	**1.12**	**1.19**	**1.10**	**1.29**	**1.11**	**1.03**	**1.21**
	NoSCAN	1.04	0.97	1.11	1.05	0.97	1.12	1.03	0.94	1.13	1.05	0.95	1.16
SIMD quintile	1	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	2	0.92	0.86	1.00	**0.90**	**0.83**	**0.97**	0.98	0.88	1.09	0.95	0.85	1.05
	3	**0.83**	**0.77**	**0.90**	**0.81**	**0.75**	**0.88**	**0.89**	**0.80**	**0.99**	**0.86**	**0.77**	**0.97**
	4	**0.73**	**0.67**	**0.79**	**0.74**	**0.68**	**0.80**	**0.74**	**0.66**	**0.83**	**0.75**	**0.66**	**0.84**
	5	**0.71**	**0.66**	**0.77**	**0.68**	**0.62**	**0.74**	**0.77**	**0.69**	**0.86**	**0.74**	**0.66**	**0.83**
Tumour Stage	I	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	II	**1.61**	**1.50**	**1.73**	**1.40**	**1.31**	**1.50**	**2.55**	**2.24**	**2.92**	**2.31**	**2.02**	**2.64**
	III	**2.33**	**2.16**	**2.52**	**2.34**	**2.17**	**2.52**	**5.61**	**4.92**	**6.39**	**5.56**	**4.88**	**6.33**
Urban/Rural	Urban	1	–	–	1	–	–	1	–	–	1	–	–
	Rural	**0.90**	**0.85**	**0.96**	0.94	0.88	1.01	**0.89**	**0.82**	**0.97**	0.96	0.87	1.05
Year of diagnosis	2008–10	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	2009–11	**0.86**	**0.80**	**0.91**	**0.93**	**0.87**	**0.99**	**0.89**	**0.81**	0.97	0.97	0.89	1.06
	2012–14	**0.74**	**0.69**	**0.79**	**0.81**	**0.75**	**0.87**	**0.74**	**0.67**	**0.82**	**0.81**	**0.73**	**0.89**
	2015–18	**0.72**	**0.66**	**0.79**	**0.77**	**0.70**	**0.84**	**0.75**	**0.67**	**0.84**	**0.78**	**0.69**	**0.87**
Inpatient episodes	Mean	**1.01**	**1.01**	**1.01**	1.00	1.00	1.01	**1.01**	**1.00**	**1.01**	1.00	1.00	1.01
Comorbidity score	Mean	1.39	0.53	1.91	**1.25**	**1.15**	**1.35**	**1.23**	**1.10**	**1.37**	**1.15**	**1.01**	**1.30**

sixfold for patients diagnosed with colon cancer and almost fourfold for those diagnosed with rectal cancer. These findings are consistent with published national statistics for colorectal cancer in the UK [29] and in other published works reporting the effect of disease stage in Cox-Proportional Hazards models of CRC patients [30–33].

The models show that in general, survival has improved consistently since 2006-08 for patients diagnosed with either colon or rectal cancer. This concurs with other evidence for Scotland and the rest of the UK [34]. Furthermore, as expected, the models find a higher comorbidity score is associated with significantly poorer OS and CRCS for patients diagnosed with both colon and rectal cancer. This finding is consistent with other evidence for the UK and the rest of the world [35–38].

Despite controlling for underlying patient comorbidity as identified from hospital records, the models suggest significant regional variation in outcomes between the three MCNs for patients diagnosed with colon cancer. In particular, patients diagnosed with colon cancer in SCAN appear to have poorer CRC outcomes compared to patients in WoSCAN, both all cause and CRC mortality specifically. These differences suggest potentially inequitable outcomes for patients in an NHS that is designed to ensure equal access to quality care. Other explanations could be selection into the early-stage surgically treated cohort, which could depend on area specific factors that determine this selection. For example, the underlying health of patients in the area; surgeon dependent influences such as comfort in resecting large tumours; and data management system differences. Supplementary Tables A, B and C display the patient characteristics before selection into the early-stage surgically treated cohort and may indicate where possible selection bias should be considered. The models also account for location via the urban/rural indicator. Previous literature shows mixed evidence on the impact of rurality on survival [8, 38] Our models support the view that rurality has no significant impact on survival. However, this may reflect the greater proportion of older individuals captured in our rural cohort compared to the general population. This finding may also differ for later stage patients, where distance

Table 5: Cox-proportional hazard regression models: patients diagnosed with early-stage rectal cancer in Scotland between January 2006 and December 2018 and treated with surgery, n = 6,440

Variable	Category	All-cause mortality						CRC-specific mortality					
		Unadjusted			Adjusted			Unadjusted			Adjusted		
		HR	95 %	CI	HR	95 %	CI	HR	95 %	CI	HR	95%	CI
Sex	Male	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	Female	**0.79**	**0.72**	**0.87**	**0.75**	**0.68**	**0.82**	**0.77**	**0.68**	**0.87**	**0.72**	**0.63**	**0.82**
Age (years)	18–34	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	35–44	1.40	0.62	3.16	1.38	0.60	3.15	1.06	0.47	2.38	1.04	0.45	2.40
	45–54	1.04	0.48	2.25	0.98	0.45	2.14	0.79	0.37	1.69	0.75	0.34	1.63
	55–64	1.27	0.60	2.71	1.20	0.56	2.60	0.92	0.44	1.93	0.88	0.41	1.89
	65–74	1.82	0.86	3.87	1.74	0.81	3.74	0.95	0.46	1.98	0.93	0.43	1.99
	75–84	**3.47**	**1.64**	**7.36**	**3.40**	**1.58**	**7.33**	1.74	0.83	3.62	1.76	0.82	3.79
	85+	**6.96**	**3.25**	**14.93**	**7.41**	**3.40**	**16.13**	**3.51**	**1.64**	**7.49**	**3.91**	**1.78**	**8.59**
Cancer Network	WoSCAN	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	SCAN	1.07	0.97	1.19	0.99	0.89	1.10	1.12	0.98	1.28	1.02	0.88	1.17
	NoSCAN	1.05	0.93	1.18	1.05	0.93	1.19	1.11	0.96	1.30	1.11	0.95	1.31
SIMD quintile	1	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	2	**0.86**	**0.75**	**0.98**	**0.85**	**0.74**	**0.98**	0.87	0.73	1.05	0.88	0.73	1.06
	3	0.87	0.76	1.01	0.91	0.79	1.06	0.91	0.76	1.10	0.96	0.79	1.16
	4	**0.81**	**0.70**	**0.93**	**0.83**	**0.72**	**0.97**	0.88	0.73	1.06	0.89	0.73	1.08
	5	**0.74**	**0.64**	**0.86**	**0.74**	**0.64**	**0.87**	**0.79**	**0.65**	**0.96**	**0.79**	**0.65**	**0.96**
Tumour Stage	I	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	II	**1.52**	**1.36**	**1.70**	**1.44**	**1.29**	**1.61**	**2.05**	**1.74**	**2.40**	**1.94**	**1.65**	**2.28**
	III	**2.00**	**1.79**	**2.24**	**2.19**	**1.96**	**2.46**	**3.36**	**2.88**	**3.92**	**3.55**	**3.04**	**4.14**
Urban/Rural	Urban	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	Remote	0.96	0.86	1.07	0.95	0.85	1.07	1.01	0.88	1.15	0.97	0.84	1.13
Year of diagnosis	2008–10	1.00	–	–	1.00	–	–	1.00	–	–	1.00	–	–
	2009–11	0.96	0.86	1.07	0.99	0.88	1.11	0.90	0.77	1.04	0.92	0.80	1.07
	2012–14	0.83	0.73	0.94	0.89	0.78	1.01	0.82	0.70	0.97	0.88	0.75	1.04
	2015–18	**0.68**	**0.57**	**0.81**	**0.75**	**0.63**	**0.90**	**0.73**	**0.59**	**0.90**	**0.79**	**0.64**	**0.98**
Inpatient episodes	Mean	**1.01**	**1.00**	**1.01**	1.00	1.00	1.01	**1.01**	**1.00**	**1.01**	1.00	1.00	1.01
Comorbidity score	Mean	**1.48**	**1.31**	**1.68**	**1.31**	**1.14**	**1.50**	**1.37**	**1.15**	**1.63**	**1.24**	**1.02**	**1.50**

to treatment centres is likely to have a bigger impact for those living in rural locations.

Another concerning finding from our study is that area-based deprivation is inversely associated with survival, even after accounting for clinical factors and patient characteristics. This finding is consistent with previous research carried out in Scotland [7]. The reasons for poorer outcomes in the most deprived areas in Scotland is potentially due to later presentation or delay in treatment, however existing evidence from one MCN finds that there is no association between deprivation and these factors [12]. Other explanations include the possibility of comorbidities present within more deprived populations that we have not been able to capture from hospital admissions data or once again the existence of inequities in access to services within these groups. Further investigation into these differences is warranted. In particular, an investigation into type, length and timing of treatment.

In addition, the models find evidence that women have a significantly lower hazard of death (i.e. better survival) compared to men. This result is consistent with previous evidence from Scotland and around the world [39, 40]. The explanation for these differences remains unclear but might be explained partly by endogenous factors such as genetics or hormones, particularly those present in younger females [40, 41]. A further explanation could be differences in access to or use of treatment; however, other evidence from the UK and the US suggests very little difference in survival between sexes [42, 43].

The main strengths of our analysis are that we have been able to utilise population level data, which provides a complete representation of the characteristics and survival outcomes of those diagnosed with CRC in Scotland. Further, the study period covers an extensive time frame which once again adds to the representation of the study population and gives external validity to the results.

However, there are some important limitations to highlight. Firstly, exploration into survival outcomes on the early-stage, surgically treated cohort may be limited by cohort selection bias and differences in data recording between MCN's, some of which can be seen in Tables A, B and C in the Appendix. Secondly, the selection of the early-stage surgically treated cohort excludes stage IV patients treated with curative intent;

therefore, the survival outcomes reported here might be higher than if those patients were included. Related to this, we have not been able to confirm the reliability of use of the therapy-objectives variable as an exclusion criteria to identify patients treated with non-curative or unknown intent. We have also not considered emergency presentations as a factor influencing survival, which may help in our understanding of differences between MCNs. Moreover, using historical datasets where key information is missing for many patients, such as disease stage, is a significant limitation. Although this is improving, linkage to other datasets including chemotherapeutic, radiation and prescribing datasets, will improve our ability to interpret findings on treatment choices and outcomes. In a similar vein, we have not been able to assess any improvement in SMR06 data quality over time. A recent publication by Public Health Scotland assessed the quality of the Scottish cancer registry in 2020/21 and found an accuracy of 98.85% for colorectal cancer [44]. If the ascertainment of colorectal cancer registrations has changed substantially from 2006, this could influence our results if those missed cases differ substantially from those included. Furthermore, we may not have fully captured patient comorbidity in the models because we only had access to hospital admission records and not to diagnoses made in primary care. Some existing research suggests that other measures of comorbidity are superior to the use of Charlson in the colorectal cancer population [45] and that deriving a Charlson index from administrative data sources may result in the under reporting of comorbidities compared to medical records [46]. Finally, the assumption of proportional hazards is violated in all of our models with the exception of the CRCS model for rectal cancer patients. The consequences of non-proportional hazards means that the power of tests is decreased, resulting in a higher likelihood of finding no significant differences when there are truly differences present. Further, for covariates with hazard ratios that increase over time, the relative risk may be overestimated, while for covariates with converging hazards, the relative risk is underestimated [28]. These consequences should therefore be considered when interpreting the model results.

Notwithstanding these limitations, we have highlighted several areas that warrant further research. Specifically, an exploration into the observed inequity in survival between MCNs for colon cancer patients to uncover whether differences in the underlying health of the population are driving variation in survival, or whether system specific factors like practitioner decision making and data management issues play a role. Key next steps in our analysis will be to investigate these areas further and utilise a number of additional administrative data sets that have been linked to the registry data for the purposes of this project [23]. These include detailed information on chemotherapy prescribing and cancer audit data.

Conclusion

In summary, we have used a nationally linked administrative dataset to retrospectively explore patterns in treatment and outcomes for CRC patients in Scotland. We have contributed to the existing evidence base, which is predominantly focused on a single geographic area for all disease stages, by providing evidence for the whole of Scotland, accounting for important prognostic factors such as disease stage. We have demonstrated that Scotland's unique data linkage infrastructure can accommodate linkage between demographic records, cancer registry and hospital admissions data, providing a fuller picture of the needs of CRC patients. We have identified a number of areas that require further research including regional and socio-demographic differences in outcomes.

Acknowledgements

This work uses data provided by patients and collected by the NHS as part of their care and support. The authors would like to acknowledge the support of the eDRIS Team (National Services Scotland) for their involvement in obtaining approvals, provisioning and linking data and the use of the secure analytical platform within the National Safe Haven. This project was funded by Cancer Research UK Scotland Centre grant CTRQQR-2021 100006

Conflict of interests

The authors have no conflict of interests to declare.

Ethics statement

Approval to access this data was granted by the Public Benefit and Privacy Panel for Health and Social Care (PBPP), project number 1718-0026. This project was approved under the favourable ethics opinion of the East of Scotland NHS Research Ethics Service for the secondary analysis of PHS data within the NSH for UK based researchers.

Funding

This project was funded by Cancer Research UK (CRUK) Ref:C23434/A23706 "Creating a UK Colorectal Cancer Intelligence Hub".

References

1. Cancer Research UK. Bowel Cancer Statistics, Accessed: 2023. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer.

2. Cancer Research UK. Bowel cancer incidence by gender and UK country, 2020. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Zero.

3. Information Services Division Scotland. Cancer incidence projections for Scotland, 2015. https://beta.isdscotland.org/find-publications-and-data/conditions-and-diseases/cancer/cancer-incidence-projections-for-scotland/.

4. Public Health Scotland. Scottish Bowel Screening Programme Statistics: For the period of invitations from May 2019 to April 2021, 2022. https://www.publichealthscotland.scot/media/11605/2022-02-08-bowel-screening-publication-report.pdf.

5. Roberta De Angelis, Milena Sant, Michel P Coleman, Silvia Francisci, Paolo Baili, Daniela Pierannunzio, Annalisa Trama, Otto Visser, Hermann Brenner, Eva Ardanaz, et al. Cancer survival in Europe 1999–2007 by country and age: results of EUROCARE5—a population-based study. *The Lancet Oncology*, 15(1):23–34, 2014. https://doi.org/10.1016/s1470-2045(13)70546-1

6. J Ferlay, M Colombet, I Soerjomataram, T Dyba, G Randi, M Bettio, A Gavin, O Visser, and F Bray. Cancer incidence and mortality patterns in Europe: estimates for 40 countries and 25 major cancers in 2018. *European Journal of Cancer*, 2018. https://doi.org/10.1016/j.ejca.2018.07.005

7. LG Shack, B Rachet, DH Brewster, and MP Coleman. Socioeconomic inequalities in cancer survival in Scotland 1986–2000. *British Journal of Cancer*, 97(7):999–1004, 2007. https://doi.org/10.1038/sj.bjc.6603980

8. NC Campbell, AM Elliott, L Sharp, LD Ritchie, J Cassidy, and J Little. Rural factors and survival from cancer: analysis of Scottish cancer registrations. *British Journal of Cancer*, 82(11):1863–1866, 2000. https://doi.org/10.1054/bjoc.1999.1079

9. G. Lyratzopoulos, J.M. Barbiere, B. Rachet, M. Baum, M.R. Thompson, and 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". *Annals of Oncology*, 22(7):1661 – 1666, 2011. http://www.sciencedirect.com/science/article/pii/S0923753419384443, https://doi.org/10.1093/annonc/mdq647

10. JJ Smith, HS Tilney, AG Heriot, AW Darzi, H Forbes, MR Thompson, JD Stamatakis, and PP Tekkis. Social deprivation and outcomes in colorectal cancer. *British Journal of Surgery*, 93(9):1123–1131, 2006. https://doi.org/10.1002/bjs.5357

11. Henrik Møller, Fredrik Sandin, David Robinson, Freddie Bray, sa Klint, Karen M Linklater, Paul C Lambert, Lars Påhlman, Lars Holmberg, and Eva Morris. Colorectal cancer survival in socioeconomic groups in England: variation is mainly in the short term after diagnosis. *European Journal of Cancer*, 48(1):46–53, 2012. https://doi.org/10.1016/j.ejca.2011.05.018

12. HM Paterson, BJ Mander, P Muir, HA Phillips, and SH Wild. Deprivation and access to treatment for colorectal cancer in southeast Scotland 2003–2009. *Colorectal Disease*, 16(2):O51–O57, 2014. https://doi.org/10.1111/codi.12442

13. David H Brewster, Catherine S Thomson, David J Hole, Roger J Black, Paul L Stroner, and Charles 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. *British Medical Journal*, 322(7290):830–831, 2001. https://doi.org/10.1136/bmj.322.7290.830

14. DJ Hole and CS McArdle. Impact of socioeconomic deprivation on outcome after surgery for colorectal cancer. *British Journal of Surgery*, 89(5):586–590, 2002. https://doi.org/10.1046/j.1365-2168.2002.02073.x

15. Public Health Scotland. Estimates of survival from colorectal cancer, 2022. Available: https://www.publichealthscotland.scot/publications/cancer-survival-statistics/cancer-survival-statistics-people-diagnosed-with-cancer-during-2015-to-2019/.

16. Elizabeth Lemmon, Catherine R Hanna, Peter Hall, and Eva Morris. Health economic studies of colorectal cancer and the contribution of administrative data: A systematic review. *European Journal of Cancer Care*, 30(5), 2021. https://doi.org/10.1111/ecc.13477

17. Roxanne Connelly, Christopher J Playford, Vernon Gayle, and Chris Dibben. The role of administrative data in the big data revolution in social science research. *Social Science Research*, 59:1–12, 2016. https://doi.org/10.1016/j.ssresearch.2016.04.015

18. Clare Pearson, Jess Fraser, Michael Peake, Roland Valori, Veronique Poirier, Victoria H Coupland, Sara Hiom, Sean McPhail, Jodie Moffat, Georgios Lyratzopoulos, et al. Establishing population-based surveillance of diagnostic timeliness using linked cancer registry and administrative data for patients with colorectal and lung cancer. *Cancer Epidemiology*, 61:111–118, 2019. https://doi.org/10.1016/j.canep.2019.05.010

19. Nicholas E Burr, Edmund Derbyshire, John Taylor, Simon Whalley, Venkataraman Subramanian, Paul J Finan, Matthew D Rutter, Roland Valori, and Eva JA Morris. Variation in post-colonoscopy colorectal cancer across colonoscopy providers in English National Health Service: population based cohort study. *British Medical Journal*, 367, 2019. https://doi.org/10.1136/bmj.l6090

20. EJA Morris, D Forman, JD Thomas, P Quirke, EF Taylor, L Fairley, B Cottier, and G Poston. Surgical management and outcomes of colorectal cancer liver metastases. *British Journal of Surgery*, 97(7):1110–1118, 2010. https://doi.org/10.1002/bjs.7032

21. John C Taylor, Daniel Swinson, Jenny F Seligmann, Rebecca J Birch, Alice Dewdney, Victoria Brown, Joanna Dent, Hannah L Rossington, Philip Quirke, Eva JA Morris, et al. Addressing the variation in adjuvant chemotherapy treatment for colorectal cancer: Can a regional intervention promote national change? *International Journal of Cancer*, 148(4):845–856, 2021. https://doi.org/10.1002/ijc.33261

22. Rebecca J Birch, John C Taylor, Amy Downing, Katie Spencer, Paul J Finan, Riccardo A Audisio, Christopher

M Carrigan, Peter J Selby, and Eva JA Morris. Rectal cancer in old age–is it appropriately managed? Evidence from population-based analysis of routine data across the English national health service. *European Journal of Surgical Oncology*, 45(7):1196–1204, 2019. https://doi.org/10.1016/j.ejso.2019.01.005

23. R Robertson, Neil Crawford Campbell, S Smith, PT Donnan, F Sullivan, R Duffy, Lewis Duthie Ritchie, D Millar, J Cassidy, and Alasdair Munro. Factors influencing time from presentation to treatment of colorectal and breast cancer in urban and rural areas. *British Journal of Cancer*, 90(8):1479–1485, 2004. https://doi.org/10.1038/sj.bjc.6601753

24. Catherine R Hanna, Elizabeth Lemmon, Holly Ennis, Robert J Jones, Joy Hay, Roger Halliday, Steve Clark, Eva Morris, and Peter Hall. Creation of the first national linked colorectal cancer dataset in Scotland: prospects for future research and a reflection on lessons learned. *International Journal of Population Data Science*, 6(1), 2021. https://doi.org/ 10.23889/ijpds.v6i1.1654

25. Information Services Division Scotland. eDRIS Frequently Asked Questions, D4: Do I need ethical approval? , 2020. https://www.isdscotland.org/Products-and-Services/eDRIS/FAQ-eDRIS/index.asp#d4.

26. Eric I Benchimol, Liam Smeeth, Astrid Guttmann, Katie Harron, David Moher, Irene Petersen, Henrik T Sørensen, Erik von Elm, Sinéad M Langan, and RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Medicine*, 12(10):e1001885, 2015. https://doi.org/10.1371/journal.pmed.1001885

27. Hude Quan, Vijaya Sundararajan, Patricia Halfon, Andrew Fong, Bernard Burnand, JeanChristophe Luthi, L Duncan Saunders, Cynthia A Beck, Thomas E Feasby, and William A Ghali. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Medical Care*, pages 1130–1139, 2005. https://doi.org/10.1097/01.mlr.0000182534.19832.83

28. Michael Schemper. Cox analysis of survival data with non-proportional hazard functions. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 41(4):455–465, 1992. https://doi.org/10.2307/2349009

29. Cancer Research UK. Bowel cancer survival by stage at diagnosis, 2019. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival#heading-Three

30. Han-Ching Chan, Chi-Cheng Huang, Ching-Chieh Huang, Amrita Chattopadhyay, KuanHung Yeh, Wen-Chung Lee, Chun-Ju Chiang, Hsin-Ying Lee, Skye Hung-Chun Cheng, and Tzu-Pin Lu. Predicting colon cancer-specific survival for the Asian population using national cancer registry data from Taiwan. *Annals of Surgical Oncology*, 29(2):853–863, 2022. https://doi.org/10.1245/s10434-021-10646-2

31. Cassia B Wang, Faisal Shahjehan, Amit Merchea, Zhuo Li, Tanios S Bekaii-Saab, Axel Grothey, Dorin T Colibaseanu, and Pashtoon M Kasi. Impact of tumor location and variables associated with overall survival in patients with colorectal cancer: A mayo clinic colon and rectal cancer registry study. *Frontiers in Oncology*, 9:76, 2019. https://doi.org/ 10.3389/fonc.2019.00076

32. Yajun Yu, Megan Carey, William Pollett, Jane Green, Elizabeth Dicks, Patrick Parfrey, Yildiz E Yilmaz, and Sevtap Savas. The long-term survival characteristics of a cohort of colorectal cancer patients and baseline variables associated with survival outcomes with or without time-varying effects. *BMC Medicine*, 17(1):1–12, 2019. https://doi.org/ 10.1186/s12916-019-1379-5

33. Lekun Fang, Ziqing Yang, Mingliang Zhang, Manqi Meng, Junyan Feng, and Chuangqi Chen. Clinical characteristics and survival analysis of colorectal cancer in China: a retrospective cohort study with 13,328 patients from southern China. *Gastroenterology Report*, 9(6):571–582, 2021. https://doi.org/10.1093/gastro/goab048

34. Melina Arnold, Mark J Rutherford, Aude Bardot, Jacques Ferlay, Therese ML Andersson, Tor Åge Myklebust, Hanna Tervonen, Vicky Thursfield, David Ransom, Lorraine Shack, et al. Progress in cancer survival, mortality, and incidence in seven high-income countries 1995–2014 (ICBP SURVMARK-2): a population-based study. *The Lancet Oncology*, 20(11):1493–1505, 2019. https://doi.org/10.1016/s1470-2045(19)30456-5

35. Maleshwane Lettie Pule, Elizabeth Buckley, Theophile Niyonsenga, and David Roder. The effects of comorbidity on colorectal cancer mortality in an Australian cancer population. *Scientific Reports*, 9(1):1–10, 2019. https://doi.org/10.1038/s41598-019-44969-8

36. Lene H Iversen, Mette Nørgaard, Jacob Jacobsen, Søren Laurberg, and Henrik T Sørensen. The impact of comorbidity on survival of Danish colorectal cancer patients from 1995 to 2006-a population-based cohort study. *Diseases of the Colon & Rectum*, 52(1):71–78, 2009. https://doi.org/10.1007/dcr.0b013e3181974384

37. Ralf Rieker, Eva Hammer, Roland Eisele, Eberhardt Schmid, and Josef Högel. The impact of comorbidity on the overall survival and the cause of death in patients after colorectal cancer resection. *Langenbeck's Archives of Surgery*, 387(2):72–76, 2002. https://doi.org/10.1007/s00423-002-0291-0

38. Colleen A Cuthbert, Brenda R Hemmelgarn, Yuan Xu, and Winson Y Cheung. The effect of comorbidities on outcomes in colorectal cancer survivors: a population-based cohort study. *Journal of Cancer Survivorship*, 12(6):733–743, 2018. https://doi.org/10.1007/ s11764-018-0710-z

39. Campbell NC, Elliott AM, Sharp L, Ritchie LD, Cassidy J, Little J. Impact of deprivation and rural residence on treatment of colorectal and lung cancer. British journal of cancer. 2002 Sep;87(6):585-90.

40. CS McArdle, DC McMillan, and DJ Hole. Male gender adversely affects survival following surgery for colorectal cancer. *British Journal of Surgery*, 90(6):711–715, 2003. https://doi.org/10.1002/bjs.4098

41. Yafan Yang, Guiying Wang, Jingli He, Shuguang Ren, Fengpeng Wu, Jianfeng Zhang, and Feifei Wang. Gender differences in colorectal cancer survival: a meta-analysis. *International Journal of Cancer*, 141(10):1942–1949, 2017. https://doi.org/10.1002/ijc.30827

42. Ondrej Majek, Adam Gondos, Lina Jansen, Katharina Emrich, Bernd Holleczek, Alexander Katalinic, Alice Nennecke, Andrea Eberle, Hermann Brenner, GEKID Cancer Survival Working Group, et al. Sex differences in colorectal cancer survival: population-based analysis of 164,996 colorectal cancer patients in Germany. *PloS One*, 8(7):e68077, 2013. https://doi.org/10.1371/journal.pone.0068077

43. Alan White, Lucy Ironmonger, Robert JC Steele, Nick Ormiston-Smith, Carina Crawford, and Amanda Seims. A review of sex-related differences in colorectal cancer incidence, screening uptake, routes to diagnosis, cancer stage and survival in the UK. *BMC Cancer*, 18(1):1–11, 2018. https://doi.org/10.1186/s12885-018-4786-7

44. Public Health Scotland. Assessment of the National Cancer Registry, Data Quality Assurance, 2022.

45. Jessica R Lieffers, Vickie E Baracos, Marcy Winget, and Konrad Fassbender. A comparison of Charlson and Elixhauser comorbidity measures to predict colorectal cancer survival using administrative health data. *Cancer*, 117(9):1957–1965, 2011. https://doi.org/10.1002/cncr.25653

46. Hude Quan, Gerry A Parsons, and William A Ghali. Validity of information on comorbidity derived from ICD-9-CCM administrative data. *Medical Care*, pages 675–685, 2002. https://doi.org/10.1097/00005650-200208000-00007

Abbreviations

AJCC:	American Joint Committee on Cancer
CRC:	Colorectal cancer
CRCS:	Colorectal cancer survival
ICD-10:	International Disease Classification 10th Revision
KM:	Kaplan-Meier
MCN:	Managed Clinical Network
NRS:	National Records of Scotland
NoSCAN:	North of Scotland Cancer Network
PHS:	Public Health Scotland's
PBPP:	Public Benefit and Privacy Panel
RECORD:	REporting of studies Conducted using Observational Routinely collected health Data
SCAN:	South of Scotland Cancer Network
SIMD:	Scottish Index of Multiple Deprivation
SMR06:	Scottish Morbidity Record 06
SMR01:	Scottish Morbidity Record 01
UK:	United Kingdom
WoSCAN:	West of Scotland Cancer Network

Supplementary Appendix 1

Table A: Patient Characteristics by Managed Cancer Network for patients diagnosed with colon or rectal cancer in Scotland between January 2006 and December 2018: Demographics

		Total cohort N = 44,875											
		Colon/Sigmoid						Rectal					
		SCAN		WoSCAN		NoSCAN		SCAN		WoSCAN		NoSCAN	
		N = 11,344		N = 13,186		N = 8,161		N = 4,110		N = 4,880		N = 3,194	
		n	%	n	%	n	%	n	%	n	%	n	%
Sex	Male	5,890	51.9	6,888	52.2	4,211	51.6	2,544	61.9	2,998	61.4	1,990	62.3
	Female	5,454	48.1	6,298	47.8	3,950	48.4	1,566	38.1	1,882	38.6	1,204	37.7
Age	18–34	112	1.0	85	0.6	68	0.8	33	0.8	39	0.8	19	0.6
	35–44	203	1.8	280	2.1	115	1.4	101	2.5	141	2.9	83	2.6
	45–54	775	6.8	997	7.6	588	7.2	419	10.2	493	10.1	315	9.9
	55–64	1,925	17.0	2,371	18.0	1,426	17.5	889	21.6	1,122	23.0	743	23.3
	65–74	3,382	29.8	4,077	30.9	2,464	30.2	1,276	31.0	1,542	31.6	958	30.0
	75–84	3,488	30.7	3,936	29.8	2,523	30.9	1,025	24.9	1,158	23.7	814	25.5
	85+	1,459	12.9	1,440	10.9	977	12.0	367	8.9	385	7.9	262	8.2
SIMD Quintile	1	1,743	15.4	3,969	30.1	584	7.2	638	15.5	1,500	30.7	230	7.2
	2	2,542	22.4	3,182	24.1	1,269	15.5	898	21.8	1,182	24.2	554	17.3
	3	2,687	23.7	2,154	16.3	2,162	26.5	905	22.0	788	16.1	823	25.8
	4	2,048	18.1	1,762	13.4	2,436	29.8	812	19.8	650	13.3	974	30.5
	5	2,324	20.5	2,119	16.1	1,710	21.0	857	20.9	760	15.6	613	19.2
Urban/Rural	Urban	8,716	76.8	11,715	88.8	4,588	56.2	3,157	76.8	4,322	88.6	1,763	55.2
	Remote or Rural	2,628	23.2	1,471	11.2	3,573	43.8	953	23.2	558	11.4	1,431	44.8
Cancer Network	SCAN	11,344	100.0	–	–	–	–	4,110	100.0	–	–	–	–
	WoSCAN	–	–	13,186	100.0	–	–	–	–	4,880	100.0	–	–
	NoSCAN	–	–	–	–	8,161	100.0	–	–	–	–	3,194	100.0

Table B: Patient Characteristics by Managed Cancer Network for patients diagnosed with colon or rectal cancer in Scotland between January 2006 and December 2018: Diagnosis

		Total cohort N = 44,875											
		Colon/Sigmoid						Rectal					
		SCAN		WoSCAN		NoSCAN		SCAN		WoSCAN		NoSCAN	
		N = 11,344		N = 13,186		N = 8,161		N = 4,110		N = 4,880		N = 3,194	
		n	%	n	%	n	%	n	%	n	%	n	%
Tumour stage	I	1,430	12.6	1,785	13.5	924	11.3	947	23.0	991	20.3	713	22.3
	II	3,231	28.5	3,310	25.1	2,368	29.0	792	19.3	869	17.8	649	20.3
	III	2,895	25.5	2,988	22.7	2,152	26.4	926	22.5	908	18.6	690	21.6
	IV	2,414	21.3	3,049	23.1	1,706	20.9	642	15.6	857	17.6	448	14.0
	Unknown	1,374	12.1	2,054	15.6	1,011	12.4	803	19.5	1,255	25.7	694	21.7
Method 1st detection	Screening	1,366	12.0	1,812	13.7	1,254	15.4	560	13.6	735	15.1	504	15.8
	Incidental finding	337	3.0	343	2.6	304	3.7	74	1.8	73	1.5	59	1.8
	Clinical presentation	9,543	84.1	10,890	82.6	6,478	79.4	3,424	83.3	4,031	82.6	2,610	81.7
	Interval cancer, other	39	0.3	23	0.2	90	1.1	18	0.4	41	0.8	21	0.7
	Unknown	59	0.5	118	0.9	35	0.4	34	0.8	82	0.7	82	0.7
Year of diagnosis	2006-2008	2,565	22.6	3,019	22.9	1,968	24.1	977	23.8	1,205	24.7	759	23.8
	2009–2011	2,881	25.4	3,380	25.6	2,037	25.0	1,105	26.9	1,185	24.3	804	25.2
	2012–2014	2,826	24.9	3,278	24.9	1,882	23.1	962	23.4	1,130	23.2	731	22.9
	2015–2018	3,072	27.1	3,509	26.6	2,274	27.9	1,066	25.9	1,360	27.9	900	28.2

Table C: Patient Characteristics by Managed Cancer Network for patients diagnosed with colon or rectal cancer in Scotland between January 2006 and December 2018:Treatment

		Total cohort N = 44,875											
		Colon/Sigmoid						**Rectal**					
		SCAN		**WoSCAN**		**NoSCAN**		**SCAN**		**WoSCAN**		**NoSCAN**	
		N = 11,344		**N = 13,186**		**N = 8,161**		**N = 4,110**		**N = 4,880**		**N = 3,194**	
		n	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Therapy objectives	Curative	7,302	64.4	7,476	56.7	4,535	55.6	2,819	68.6	2,836	58.1	1,930	60.4
	Palliative	3,783	33.3	4,928	37.4	2,691	33.0	1,173	28.5	1,659	34.0	827	25.9
	Unknown	259	2.3	782	5.9	935	11.5	118	2.9	385	7.9	437	13.7
Chemotherapy	No	8,093	71.3	8,991	68.2	5,639	69.1	2,543	61.9	2,534	51.9	1,845	57.8
	Yes/planned	3,192	28.1	3,950	30.0	2,512	30.8	1,544	37.6	2,248	46.1	1,344	42.1
	Not known	59	0.5	245	1.9	10	0.1	23	0.6	98	2.0	5	0.2
Radiotherapy	No	11,060	97.5	12,635	95.8	7,861	96.3	2,534	61.7	2,917	59.8	1,754	54.9
	Yes/planned	218	1.9	257	1.9	300	3.7	1,539	37.4	1,818	37.3	1,436	45.0
	Not known	66	0.6	294	2.2	368	1.1	37	0.9	1	0.0	4	0.1
Surgery	No	2,516	22.2	3,716	28.2	1,914	23.5	991	24.1	1,502	30.8	867	27.1
	Yes/planned	8,807	77.6	9,386	71.2	6,247	76.5	3,108	75.6	3,331	68.3	2,326	72.8
	Not known	21	0.2	84	0.6	110	0.3	11	0.3	47	1.0	1	0.0
Inpatient episodes	Mean	1.95	–	2.20	–	1.75	–	1.99	–	2.07	–	1.77	–
Comorbidity score	Mean	0.07	–	0.07	–	0.07	–	0.06	–	0.05	–	0.07	–