Factors associated with advanced colorectal cancer differ between young and older adults in England: a population-based cohort study

A. C. Chambers*'†, S. W. Dixon†, P. White‡, M. G. Thomas*, A. C. Williams† and D. E. Messenger* iD

*Department of Colorectal Surgery, University Hospitals Bristol, Bristol Royal Infirmary, Bristol, UK, †School of Cellular and Molecular Medicine, University Walk, University of Bristol, Bristol, UK, and ‡Department of Engineering Design and Mathematics, University of West of England, Bristol, UK

Abstract

Aim Advanced stage presentation of colorectal cancer is associated with poorer survival outcomes, particularly among young adults. This study aimed to determine whether demographic risk factors for advanced stage presentation differed between young and older adults.

Method Individual-level data on all incident colorectal cancers in people aged 20 years and above were extracted from the National Cancer Registration and Analysis Service database for the years 2012 to 2015. Patients were divided into two cohorts: young-onset colorectal cancer (YOCC) if aged 20–49 years and older-onset colorectal cancer (OOCC) if aged 50 years and above. Logistic regression was used to identify risk factors for advanced stage presentation, defined as TNM Stage III or IV, in each cohort.

Results There were 7075 (5.2%) patients in the YOCC cohort and 128 345 (94.8%) patients in the OOCC cohort. Tumours in the YOCC cohort were more likely to be at an advanced stage (67.2% *vs* 55.3%, *P* < 0.001) and located distally (63.7% *vs* 55.4%, *P* < 0.001). No demographic factor was consistently associated with advanced stage presentation in the YOCC cohort. Among the OOCC cohort, increased social deprivation [OR (Index of Multiple Deprivation quintile 5 *vs* 1) = 1.11 (95% CI 1.07–1.16), *P* < 0.001], Black/ Black British ethnicity [OR (baseline White) = 1.25 (95% CI 1.11–1.40), *P* < 0.001] and residence in the East Midlands [OR (baseline London) = 1.11 (95% CI 1.04–1.17), *P* = 0.001] were associated with advanced stage presentation.

Conclusion Demographic factors associated with advanced disease were influenced by age. The effects of social deprivation and ethnicity were only observed in older adults and mirror trends in screening uptake. Targeted interventions for high-risk groups are warranted.

Keywords Colorectal neoplasms, demography, risk factors, age distribution, socioeconomic factors, ethnic groups

What does this paper add to the literature?
Demographic risk factors for advanced stage presentation of colorectal cancer are not well described in the UK. Although younger age was associated with more advanced and distal disease, inequalities related to socioeconomic status and ethnicity were only noted in older adults and reflect similar trends in screening uptake.

Introduction

In England, colorectal cancer (CRC) is the third most common cancer in men and women and accounts for 35 000 new diagnoses each year [1]. While overall age-standardized incidence rates of CRC have remained stable in the UK [2,3], this trend masks recent changes in age-specific incidence rates. There is a substantial body of evidence to suggest that the incidence of CRC in adults under 50 years is rapidly increasing in England and other nations with a high human development index (HDI) [4–9], although the cause of this increase is not well understood and is thought to be associated with a rising prevalence of obesity, adoption of sedentary lifestyles and changes to the gut microbiome [10–13]. Conversely, recent studies from the United States have shown a decline in incidence rates in adults aged over 50 years that has largely been attributed to the

introduction of screening [5,14,15].There is emerging evidence that young-onset and older-onset CRC may in fact be biologically distinct entities. Despite a higher prevalence of hereditary syndromes in young-onset CRC, most tumours are sporadic in nature and typically located distally [16], display microsatellite stability and are CpG island methylator phenotype-low [17–19].

Advanced stage presentation is associated with worse survival outcomes, with 5-year net survival in England ranging from 10.3% in Stage IV disease to 91.7% in Stage I disease [20]. Population-based studies in other nations with a high HDI have shown that younger age, social deprivation and Black ethnicity are associated with advanced stage CRC [21–25], but demographic risk factors have not been explored in detail and it is unknown whether these risk factors differ between young and older adults. In the UK, there is scant evidence regarding the association between demographic factors, such as socioeconomic status (SES) and ethnicity, and advanced stage presentation due to a lack of complete data in nationally curated cancer registries until recently [26]. Take-up of the Bowel Cancer Screening Programme (BCSP) in England is known to be reduced in men, more deprived socioeconomic groups and areas of increased ethnic diversity [27], but how these trends are reflected in the stage at presentation of the screening age population has yet to be demonstrated. Identifying the demographic factors associated with advanced stage presentation is of critical importance as it would help guide targeted interventions at a population level.

Therefore, the aim of this study was to determine whether demographic risk factors for advanced stage presentation of CRC differed between young and older adults in the English population using data from the National Cancer Registration and Analysis Service (NCRAS) from 2012 to 2015 inclusive.

Method

This study was reported according to the STROBE guidelines for epidemiological studies (Table S1 in the online Supporting Information) [28]. Data were obtained on all patients aged 20 years and above from 2012 to 2015 using data from the National Cancer Registration and Analysis Service (NCRAS) (request ID ODR1718_067). NCRAS is the population-based cancer registry for England, operated by Public Health England, which collects, quality assures and analyses data on all patients with cancer in the English population. NCRAS receives data from multiple sources across the National Health Service including multidisciplinary team (MDT) meetings, pathology reports, molecular

testing results, treatment records, hospital activity records, hospital patient administration systems and operational standards, such as cancer waiting times, to ensure complete coverage [26].

Procedures

All new diagnoses of CRC were identified using the ICD10 codes for colonic (C18.1, C18.2-C18.9), rectosigmoid (C19) and rectal (C20) tumours. Appendiceal tumours (C18.1) were excluded due to their differing aetiology. Patients diagnosed prior to 2012 were not included because of the high level of incompleteness for stage and ethnicity data.

Patients diagnosed when under 50 years of age comprised the young-onset colorectal cancer (YOCC) cohort and patients diagnosed when aged 50 years or above comprised the older-onset colorectal cancer (OOCC) cohort. Patients were further stratified according to gender, anatomical subsite (proximal – caecum to descending colon; distal – sigmoid to rectum), geographical region, Index of Multiple Deprivation (IMD) quintile and ethnic group. The same anatomical subsite definition was used as in a previous study by the authors to ensure consistency [4]. The IMD is an area-based metric that combines weighted information on the following seven domains: income, employment, education, health, crime, barriers to housing and services and the living environment. England is divided into 32 844 lower-layer super output areas (LSOAs), each of which is ascribed a score based on these domains. The LSOAs are then ranked from least to most deprived and divided into five equal groups or quintiles (1, least deprived; 5, most deprived). Ethnicity was categorized according to the harmonized ethnic group categories for England used by the Office of National Statistics [29].

Tumours were staged as Union for International Cancer Control Stage I to IV based on the TNM classification system. Coding changes between subsequent editions of the TNM system affected sub-classifications within individual stages (e.g. Stage IIIA, IIIB), but not the broader stage categorization used in this study. Staging information inputted into the NCRAS registry was obtained from MDTs, pathology reports, imaging results and post mortems to determine a single anatomical stage at diagnosis [26]. The primary outcome of advanced stage presentation was defined as Stage III or IV disease at diagnosis.

Statistical analysis

Categorical data were presented as frequencies and percentages. Unadjusted comparisons between

variables were made using the chi-square test for independence and the chi-square test for trend as appropriate. Logistic regression models based on complete case analysis were constructed with stage as the dependent variable, dichotomized as early (Stage I/II) or advanced (Stage III/IV), for the YOCC and OOCC cohorts. Given the exploratory nature of the study in determining potential risk factors for advanced stage presentation, all predictor variables were included in the regression models. The potential modifying effect of SES on the association between ethnicity and stage at presentation was explored using the chi-square test for heterogeneity of ORs for unadjusted estimates and the likelihood ratio test for interaction between ethnic group and IMD quintile for adjusted estimates derived by logistic regression. The robustness of the findings from the main analysis was assessed using sensitivity analyses with different definitions of advanced stage disease, i.e. Stage I/II/III versus IV or Stage I versus Stage II/III/IV. Potential bias arising from missing stage information was assessed using multiple imputation under the assumption that stage data were missing at random. The same predictor variables were included in the imputation models as in the original regression analyses using 20 imputed data sets. All analyses were conducted using STATA 13.0 (StataCorp. 2014, Stata Statistical Software, Release 13, StataCorp LP, College Station, Texas, USA).

Results

The YOCC cohort accounted for 7075 of 135 420 (5.2%) new cases of CRC diagnosed between 2012 and 2015 in adults aged over 20 years. Data were 100% complete for each of the predictor variables. Stage data were missing in 715 of 7075 patients (10.1%) in the YOCC cohort and in 18 234 of 128 345 patients (14.2%) in the OOCC cohort. Patients in the YOCC cohort were more likely to be female (47.5% vs 43.9%, $P < 0.001$), of non-White ethnicity (18.3% vs 8.3%, $P < 0.001$), less affluent (IMD quintiles 1 and 2: 39.6% vs 45.5%, $P < 0.001$) and to live in London (15.6% vs 8.8%, $P < 0.001$) than patients in the OOCC cohort (Table 1).The YOCC cohort was more likely than the OOCC cohort to present with advanced stage disease (60.5% vs 47.4%, $P < 0.001$) and with tumours located in the distal colorectum (63.7% vs 55.4%, $P < 0.001$). The proportion of advanced stage presentations reduced with increasing age, particularly among patients aged 60–79 years, for both proximal ($P < 0.001$) and distal tumours ($P < 0.001$; Figs 1 and 2).

Risk factors for advanced stage presentation in the YOCC cohort

Among the YOCC cohort, other/not specified ethnicity (OR = 0.81, 95% CI 0.66–1.00, $P = 0.045$) compared with White ethnicity and residence in the East of England (OR = 0.76, 95% CI 0.62–0.94, $P = 0.010$) compared with London were independently associated with reduced likelihood of advanced stage presentation in the main analysis (Table 2). The effect of ethnicity on advanced stage presentation was consistent across all IMD quintiles for each ethnic group (Table S2). Multiple imputation for missing stage data revealed minimal differences in effect estimates for any of the predictor variables compared with the main analysis (Table S3). However, none of the predictor variables were consistently associated with advanced stage presentation when using alternative cut-offs for this outcome in the sensitivity analysis (Tables S4 and S5).

Risk factors for advanced stage presentation in the OOCC cohort

Advanced stage presentation was associated with Black/Black British (OR = 1.25, 95% CI 1.11–1.40, p < 0.001) or other/not specified ethnicity (OR = 1.07, 95% CI 1.01–1.14, $P = 0.015$) compared with White ethnicity, increasing levels of social deprivation (OR (IMD quintile 5 vs 1) = 1.11, 95% CI 1.07–1.16, $P < 0.001$] and residence in the East Midlands (OR = 1.11, 95% CI 1.04–1.17, $P = 0.001$) compared with London in the main analysis (Table 3). Compared with patients aged 50–59 years, patients in each of the age categories between 60 and 89 years were less likely to present with advanced stage disease (all $P < 0.001$). The likelihood of advanced stage presentation was also reduced among patients living in Yorkshire and The Humber (OR = 0.94, 95% CI 0.89–1.00, $P = 0.042$), the South West (OR = 0.88, 95% CI 0.84–0.93, $P < 0.001$) and the East of England (OR = 0.91, 95% CI 0.87–0.96, $P = 0.001$) compared with London, and among patients with distal tumours (OR = 0.89, 95% CI 0.87–0.92, $P < 0.001$). The effect of ethnicity on advanced stage presentation was not modified by IMD quintile for Black/Black British, Asian/Asian British and Mixed ethnicity compared with White ethnicity (Table S6). There was evidence of a deprivation gradient for other/not specified ethnicity ($P = 0.051$) in the unadjusted analysis, but this did not reach statistical significance and interaction between ethnicity and IMD quintile was not demonstrated in the adjusted analysis. Effect estimates

Table I Demographic characteristics of the young-onset and older-onset colorectal cancer cohorts.

| Characteristic	Young-onset colorectal cancer (n, %)	Older-onset colorectal cancer (n, %)	P-value
Gender			
Female	3360 (47.5)	56 373 (43.9)	< 0.001
Male	3715 (52.5)	71 972 (56.1)	
Ethnicity			
White	5778 (81.7)	117 635 (91.7)	< 0.001
Black/Black British	255 (3.6)	1517 (1.2)	
Asian/Asian British	431 (6.1)	2127 (1.7)	
Mixed	71 (1.0)	310 (0.2)	
Other/not specified	540 (7.6)	6756 (5.3)	
Index of Multiple Deprivation quintile			
1	1363 (19.3)	29 144 (22.7)	< 0.001
2	1438 (20.3)	29 276 (22.8)	
3	1418 (20.0)	27 102 (21.1)	
4	1494 (21.1)	23 687 (18.5)	
5	1362 (19.3)	19 136 (14.9)	
English region			
London	1102 (15.6)	12 602 (9.8)	< 0.001
North West	948 (13.4)	18 099 (14.1)	
Yorkshire and The Humber	609 (9.5)	12 633 (9.8)	
North East	318 (4.5)	6982 (5.4)	
West Midlands	740 (10.5)	14 161 (11.0)	
East Midlands	590 (8.3)	11 394 (8.9)	
South West	744 (10.5)	15 539 (12.1)	
East of England	781 (11.0)	15 015 (11.7)	
South East	1183 (16.7)	21 920 (17.1)	
Tumour location			
Proximal	2568 (36.3)	57 239 (44.6)	< 0.001
Distal	4507 (63.7)	71 106 (55.4)	
UICC Stage			
I/II	2083 (29.4)	49 248 (38.4)	< 0.001
III/IV	4277 (60.5)	60 863 (47.4)	
Missing	715 (10.1)	18 234 (14.2)	

remained largely unchanged in the multiple imputation model for missing stage data (Table S7), except for a reduced likelihood of advanced stage presentation in patients aged ≥ 90 years compared with 50–59 years (OR = 0.84, 95% CI 0.78–0.90, P < 0.001) and residence in the South East (OR = 0.95, 95% CI 0.90–0.99, P = 0.030) compared with London. The strength and direction of the associations between advanced stage presentation and age, level of social deprivation, geographical region and tumour location remained consistent in both the regression models in the sensitivity analysis (Tables S6 and S7). The association between Black/Black British or other/not specified ethnicity and advanced stage presentation persisted in the regression model with Stages I/II/III versus Stage IV as the outcome (Table S8), but not Stage I versus Stages II/III/IV (Table S9).

Discussion

This is the largest study based on English cancer registry data to explore risk factors associated with advanced stage presentation of any cancer type. It is unique in being the first study from the UK to demonstrate evidence of socioeconomic and ethnic inequalities in the stage at presentation of CRC. Young adults comprised only a small proportion of new diagnoses of CRC but were more likely to present with advanced disease than older adults. Furthermore, they were more likely to present with distal tumours, supporting the findings of other smaller cohort studies that there may in fact be biological differences between young and older onset disease [17–19].Of note, demographic risk factors associated with advanced stage presentation differed between the YOCC and OOCC cohorts. Among
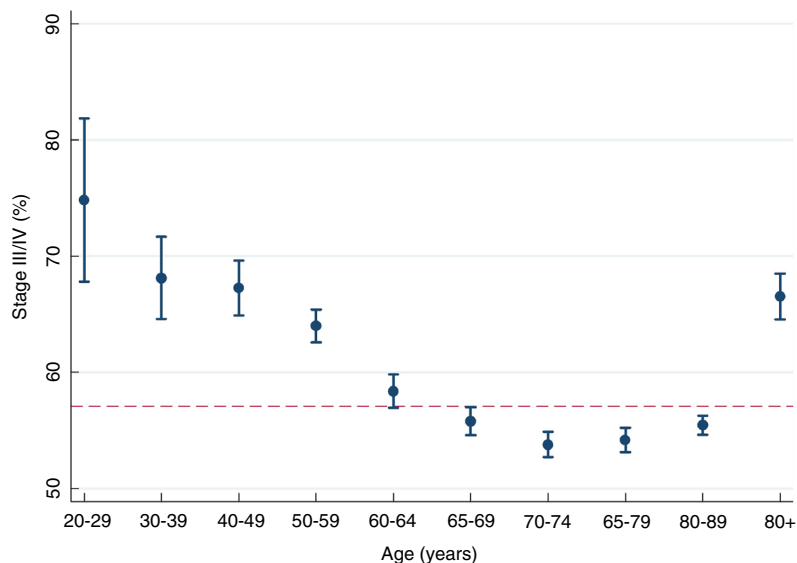
Figure 1 Proximal tumours: proportion of advanced stage presentation, defined as TNM Stage III or IV, by age category (complete case analysis). The dashed red line refers to the overall proportion of advanced stage presentations.
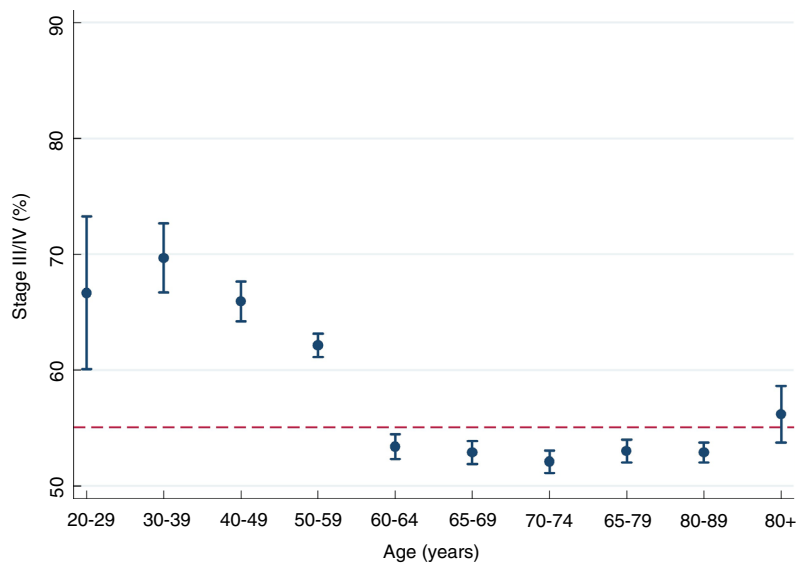
Figure 2 Distal tumours: proportion of advanced stage presentation, defined as TNM Stage III or IV, by age category (complete case analysis). The dashed red line refers to the overall proportion of advanced stage presentations.

the OOCC cohort, increasing social deprivation, Black/Black British ethnicity and residence in the East Midlands were strongly associated with advanced stage presentation, and these findings remained robust to the alternative definitions of advanced stage in the sensitivity analysis. Risk factors for advanced stage presentation were less consistent across the main and sensitivity analyses in the YOCC cohort, with no difference in risk associated with SES, ethnicity and geographical region.

The negative association between age and stage at presentation is well described, and was again demonstrated in this study for both proximal and distal tumours [21,25]. Reasons for advanced stage presentation in young adults remain unclear, despite the increased prevalence of hereditary tumours in this group [30]. Recent cohort studies suggest that while young adults experience a longer time to diagnosis than older adults this delay does not entirely explain the increased

Table 2 Logistic regression model of demographic factors associated with advanced stage presentation, defined as Stages I/II versus Stages III/IV, in the young-onset colorectal cancer cohort.

Characteristic	Early: Stages I/II (n, %)	Advanced: Stages III/IV (n, %)	Adjusted OR (95% CI)	P-value
Age (years)				
20–29	103 (29.9)	242 (70.1)	1	
30–39	489 (31.0)	1090 (69.0)	0.95 (0.73–1.22)	0.664
40–49	1491 (33.6)	2945 (66.4)	0.83 (0.65–1.06)	0.135
Gender				
Female	987 (32.5)	2053 (67.5)	1	
Male	1096 (33.0)	2224 (67.0)	0.98 (0.88–1.09)	0.678
Ethnicity				
White	1700 (32.3)	3556 (67.7)	1	
Black/Black British	80 (34.8)	150 (65.2)	0.82 (0.62–1.10)	0.189
Asian/Asian British	129 (33.7)	254 (66.3)	0.86 (0.69–1.08)	0.195
Mixed	17 (27.9)	44 (72.1)	1.16 (0.66–2.03)	0.617
Other/not specified	157 (36.5)	273 (63.5)	0.81 (0.66–1.00)	0.045
Index of Multiple Deprivation quintile				
1	393 (33.2)	790 (66.8)	1	
2	429 (33.0)	872 (67.0)	1.01 (0.85–1.19)	0.945
3	431 (33.9)	842 (66.1)	0.97 (0.82–1.14)	0.692
4	445 (32.6)	922 (67.4)	1.01 (0.86–1.20)	0.884
5	385 (31.2)	851 (68.8)	1.08 (0.90–1.29)	0.413
English region				
London	302 (31.2)	665 (68.8)	1	
North West	275 (31.9)	587 (68.1)	0.94 (0.76–1.15)	0.540
Yorkshire and The Humber	207 (33.1)	418 (66.9)	0.89 (0.71–1.11)	0.315
North East	99 (33.3)	208 (66.7)	0.92 (0.70–1.23)	0.589
West Midlands	193 (29.1)	470 (70.9)	1.09 (0.87–1.36)	0.454
East Midlands	162 (31.8)	347 (68.2)	0.96 (0.75–1.21)	0.716
South West	236 (34.6)	446 (65.6)	0.83 (0.67–1.03)	0.095
East of England	274 (37.0)	467 (63.0)	0.76 (0.62–0.94)	0.010
South East	335 (33.4)	669 (66.6)	0.89 (0.73–1.08)	0.235
Tumour location				
Proximal	749 (32.0)	1590 (68.0)	1	
Distal	1334 (33.2)	2687 (66.8)	0.96 (0.86–1.07)	0.437

risk of advanced stage presentation and that biological factors may play a role [31]. Among the OOCC cohort, adults of screening age were the least likely to present with advanced disease, which may indicate the effect of screening on this population. This would support the findings from the initial randomized controlled trials of faecal occult blood testing that demonstrated a shift to earlier stage disease at presentation among participants in the screening arms [32,33].

The deprivation gradient associated with advanced stage presentation has been reported in other HDI nations [22,23], although this relationship was not demonstrated by previous studies conducted in the UK [34,35]. In this study, there was clear evidence of a deprivation gradient among the OOCC cohort but not the YOCC cohort, implying that age is an effect modifier in the relationship between SES and stage. Studies

where analyses have been stratified by age report conflicting results. A population-based study from Florida demonstrated a deprivation gradient in both young and older adult cohorts [36]; whereas a New Hampshire study identified college-level education to be a risk factor for advanced stage presentation in young adults, but protective in older adults [37]. In the current study, the deprivation gradient among the OOCC cohort mirrors the deprivation gradient observed in screening uptake in England [27]. This may explain the lack of association between SES and stage in previous studies from the UK, as these were based on data collected prior to, or during the initial phase of, the screening programmes in England and Scotland [34,35]. Despite only small differences in the absolute risk of advanced stage presentation across IMD quintiles, elimination of the deprivation gradient would have resulted in the diagnosis of about

Table 3 Logistic regression model of demographic factors associated with advanced stage presentation, defined as Stages I/II versus Stages III/IV, in the older-onset colorectal cancer cohort.

Characteristic	Early: Stage I/II (*n*, %)	Advanced: Stage III/IV (*n*, %)	Adjusted OR (95% CI)	*P*-value
Age (years)				
50–59	5037 (37.3)	8477 (62.7)	1	
60–64	5784 (44.9)	7111 (55.1)	0.73 (0.70–0.77)	< 0.001
65–69	7623 (46.0)	8960 (54.0)	0.70 (0.67–0.73)	< 0.001
70–74	8598 (47.2)	9631 (52.8)	0.66 (0.63–0.69)	< 0.001
75–79	8610 (46.5)	9920 (53.5)	0.68 (0.65–0.71)	< 0.001
80–89	12 159 (45.8)	14 398 (54.2)	0.70 (0.68–0.73)	< 0.001
90+	1437 (37.8)	2366 (62.2)	0.97 (0.90–1.04)	0.385
Gender				
Female	21 207 (44.7)	26 253 (55.3)	1	
Male	28 041 (44.8)	34 610 (55.2)	1.02 (0.99–1.04)	0.179
Ethnicity				
White	45 670 (44.9)	55 960 (55.1)	1	
Black/Black British	495 (36.9)	845 (63.1)	1.25 (1.11–1.40)	< 0.001
Asian/Asian British	816 (43.9)	1044 (56.1)	0.97 (0.88–1.06)	0.489
Mixed	123 (44.1)	156 (55.9)	0.98 (0.77–1.25)	0.888
Other/not specified	2144 (42.9)	2858 (57.1)	1.07 (1.01–1.14)	0.015
Index of Multiple Deprivation quintile				
1	11 528 (45.9)	13 569 (54.1)	1	
2	11 460 (45.5)	13 756 (54.6)	1.02 (0.99–1.06)	0.259
3	10 374 (44.7)	12 814 (55.3)	1.05 (1.01–1.09)	0.009
4	8919 (44.1)	11 317 (55.9)	1.07 (1.03–1.11)	0.001
5	6967 (42.6)	9407 (57.4)	1.11 (1.07–1.16)	< 0.001
English region				
London	4582 (43.1)	6050 (56.9)	1	
North West	6939 (43.6)	8984 (56.4)	1.02 (0.97–1.07)	0.484
Yorkshire and The Humber	5042 (45.6)	6028 (54.4)	0.94 (0.89–1.00)	0.042
North East	2838 (45.1)	3455 (54.9)	0.95 (0.89–1.01)	0.102
West Midlands	5274 (42.9)	7029 (57.1)	1.05 (0.90–1.10)	0.097
East Midlands	3767 (41.8)	5249 (58.2)	1.11 (1.04–1.17)	0.001
South West	6487 (47.4)	7208 (52.6)	0.88 (0.84–0.93)	< 0.001
East of England	6285 (46.5)	7222 (53.5)	0.91 (0.87–0.96)	0.001
South East	8034 (45.5)	9638 (54.5)	0.96 (0.91–1.01)	0.078
Tumour location				
Proximal	21 008 (43.5)	27 334 (56.5)	1	
Distal	28 240 (45.7)	33 529 (54.3)	0.89 (0.87–0.92)	< 0.001

320 fewer cases of advanced stage disease in the OOCC cohort annually. The absence of a deprivation gradient in the YOCC cohort is difficult to explain, although it does reflect the findings from a recent study of demographic trends in the incidence of young-onset CRC in England that showed increases in the incidence rate in CRC to be similar across all IMD quintiles [4].

In the OOCC cohort, Black/Black British ethnicity was associated with a 25% increase in the odds of advanced stage presentation compared with White ethnicity. This may represent an underestimate as Black ethnicity is more frequently misclassified than White ethnicity and may explain why other/not specified ethnicity was also associated with an increase in the odds of advanced stage presentation [38]. In the USA, the association between Black ethnicity and advanced stage presentation is well described and does not appear to be modified by age [24,36]. Paradoxically, other/not specified ethnicity was shown to be protective for advanced stage presentation in the YOCC cohort. This finding cannot be readily explained and may be the result of type 1 error as the associated *P*-value was only just below the 5% threshold for statistical significance. In the UK, Black and minority ethnicity is linked to poorer health outcomes in older adults [39,40], although the effect of ethnicity on stage at

presentation for any cancer type has never been studied due to historically low levels of completeness of ethnicity data in English cancer registries [26]. The increased risk of advanced stage presentation among adults of Black/Black British ethnicity in the OOCC cohort may be attributed to screening behaviour, as screening uptake is known to be lower in areas of increased ethnic diversity [27]. However, advanced stage presentation was not associated with Asian/Asian British or Mixed ethnicity, indicating that other factors, such as lifestyle or diet, may offset the risk attributed to low screening uptake in these ethnic groups. There was no evidence that deprivation modified the effect of Black/Black British ethnicity on advanced stage presentation in either the YOCC or OOCC cohorts. In the OOCC cohort, this suggests that the increased risk of advanced stage presentation among adults of Black/Black British ethnicity could have been related to similarly low levels of screening uptake across all IMD quintiles, but also to potential differences in tumour biology compared with other ethnicities that predisposed to more rapid disease progression. Evidence for genetic differences conferring adverse outcomes in Black patients is limited to one small retrospective study and merits further investigation [41].

The geographical variation in advanced stage presentation was most pronounced in the OOCC cohort. Residence in the East Midlands was a risk factor for advanced stage presentation, whereas residence in the South West and East of England was shown to be protective. It is difficult to reconcile the increased risk associated with residence in the East Midlands with regional screening uptake rates, as first-time uptake is known to be higher in the East Midlands compared with the baseline reference region of London [27]. A recent report from the National Cancer Intelligence Network suggests that the most likely cause of geographical variation in stage at presentation for all cancers is the completeness of stage data [42]. Regions with the highest stage completeness reported the highest proportion of early stage disease, with Trent (now East Midlands) reporting the lowest proportions of stage completeness and early stage disease [42]. In the current study, it is noteworthy that the East Midlands also had the lowest stage completeness of all the English regions at 79.2%, compared with 85.8% for the entire OOCC cohort (data not shown), although stage completeness was similarly low in the South East (80.6%) with no associated increase in the risk of advanced stage presentation. Also, multiple imputation was used to minimize bias associated with missing data and the effect estimates from the multiple imputation models were essentially unchanged from

those obtained in the main analysis. Thus, variation in stage completeness by region may not have had a significant effect on stage at presentation.

The key strengths of this study are the size and completeness of the dataset, which has allowed the first national-level analysis of predictors for advanced stage presentation of CRC to be undertaken in England. However, there are some important limitations. This study was exploratory in nature and not powered to detect a prespecified difference in outcome for a given predictor variable. The YOCC cohort was large ($n = 7075$), but considerably smaller than the OOCC cohort ($n = 128\ 345$) and therefore more susceptible to type II error. Combining the data presented here with other international datasets would reduce the impact of type II error, but the benefits of doing so would have to be weighed against the potential heterogeneity introduced by including data from countries where data may be less complete, or less accurate, and healthcare systems do not provide universal coverage. Another limitation is that IMD, geographical region and ethnicity are all group-level metrics that do not fully account for individual-level contextual factors that may have influenced the association between these variables and stage at presentation. Furthermore, non-White ethnicity is more likely to be misclassified than White ethnicity and thus the effect estimates of ethnic group on stage at presentation should be viewed with caution [38].

In summary, young adults with CRC are more likely to present with advanced disease and distal tumours than older adults. Demographic factors associated with advanced stage presentation differ between young and older populations. There is clear evidence of a deprivation gradient and increased risk of advanced stage presentation associated with Black/Black British ethnicity in older adults that is not observed in the young, which reflect similar patterns in screening uptake. The rising incidence of CRC in young adults in England is characterized by a strong cohort effect with the increased risk observed in the youngest cohorts, i.e. those aged 20–29 and 30–39 years, likely to be carried forward as they age [4]. More than 60% of all tumours among adults aged 50–59 years were Stage III/IV at presentation and this age group is likely to face the dual burden of rising CRC incidence and advanced stage presentation unless the screening age is lowered. The Bowel Scope Screening Programme, which commenced in 2013, offers a one-off flexible sigmoidoscopy at the age of 55 years, but only covers half of all general practice populations in England and has the lowest uptake in the most deprived and ethnically diverse areas [43,44]. Therefore, strategies that are specifically tailored to groups at the highest risk of advanced stage presentation are warranted. These should

involve a more nuanced approach to the use of quantitative faecal immunohistochemical testing (qFIT), which has already been shown to increase screening uptake among previous nonresponders [45]. Broadening access to diagnostic endoscopy has significant resource implications in the current healthcare climate and qFIT should be used as a risk stratification tool in symptomatic younger patients.

Acknowledgements

Conflicts of interest

No conflicts of interest.

Author contributions

All authors contributed to the conception and design; acquisition of data, analysis and interpretation of data; drafting the article or revising it critically for important intellectual content; and final approval of the version to be published.

References

1 Cancer Research UK. *Bowel Cancer Statistics*. Available at http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Zero. (accessed 9 June 2020).

2 Cancer Research UK. *Bowel Cancer Statistics*. Available at http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Two. (accessed 9 June 2020).

3 Arnold M, Sierra MS, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 2017; **66:** 683–91.

4 Chambers AC, Dixon SW, White P, Williams AC, Thomas MG, Messenger DE. Demographic trends in the incidence of young-onset colorectal cancer: a population-based study. *Br J Surg* 2020; **107:** 595–605.

5 Siegel RL, Fedewa SA, Anderson WF *et al.* Colorectal cancer incidence patterns in the United States, 1974–2013. *J Natl Cancer* 2017; **109:** 1974–2013.

6 Brenner DR, Ruan Y, Shaw E, De P, Heitman SJ, Hilsden RJ. Increasing colorectal cancer incidence trends among younger adults in Canada. *Prev Med* 2017; **105:** 345–9.

7 Boyce S, Nassar N, Lee CY, Suen MK, Al Zahrani S, Gladman MA. Young-onset colorectal cancer in New South Wales: a population-based study. *Med J Aust* 2016; **205:** 465–70.

8 Gandhi J, Davidson C, Hall C *et al.* Population-based study demonstrating an increase in colorectal cancer in young patients. *Br J Surg* 2017; **104:** 1063–8.

9 Vuik FE, Nieuwenburg SA, Bardou M *et al.* Increasing incidence of colorectal cancer in young adults in Europe over the last 25 years. *Gut* 2019; **68:** 1820–6.

10 Danaei G, Vander Hoorn S, Lopez AD, Murray CJL, Ezzati M. Comparative Risk Assessment collaborating group (Cancers). Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. *Lancet* 2005; **366:** 1784–93.

11 Kantor ED, Udumyan R, Signorello LB, Giovannucci EL, Montgomery S, Fall K. Adolescent body mass index and erythrocyte sedimentation rate in relation to colorectal cancer risk. *Gut* 2016; **65:** 1289–95.

12 Liu PH, Wu K, Ng K *et al.* Association of obesity with risk of early-onset colorectal cancer among women. *JAMA Oncol* 2018; **5:** 37–44.

13 Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol* 2014; **12:** 661–72.

14 Edwards BK, Ward E, Kohler BA *et al.* Annual report to the nation on the status of cancer, 1975–2006, featuring colorectal cancer trends and impact of interventions (risk factors, screening, and treatment) to reduce future rates. *Cancer* 2010; **116:** 544–73.

15 Levin TR, Corley DA, Jensen CD *et al.* Effects of organized colorectal cancer screening on cancer incidence and mortality in a large community-based population. *Gastroenterology* 2018; **155:** 1383–91.

16 Segev L, Kalady MF, Church JM. Left-sided dominance of early-onset colorectal cancers: a rationale for screening flexible sigmoidoscopy in the young. *Dis Colon Rectum* 2018; **61:** 897–902.

17 Willauer AN, Liu Y, Pereira AAL *et al.* Clinical and molecular characterization of early-onset colorectal cancer. *Cancer* 2019; **125:** 2002–10.

18 Kirzin S, Marisa L, Guimbaud R *et al.* Sporadic early-onset colorectal cancer is a specific sub-type of cancer: a morphological, molecular and genetics study. *PLoS One* 2014; **9:** e103159.

19 Perea J, Rueda D, Canal A *et al.* Age at onset should be a major criterion for subclassification of colorectal cancer. *J Mol Diagn* 2014; **16:** 116–26.

20 Cancer Research UK. *Bowel Cancer Statistics*. Available at https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/survival#heading-Three. (accessed 9 June 2020)

21 Fairley TL, Cardinez CJ, Martin J *et al.* Colorectal cancer in U.S. adults younger than 50 years of age, 1998–2001. *Cancer* 2006; **107:** 1153–61.

22 Feller A, Schmidlin K, Bordoni A *et al.* Socioeconomic and demographic inequalities in stage at diagnosis and survival

among colorectal cancer patients: evidence from a Swiss population-based study. *Cancer Med* 2018; 7: 1498–510.

23 Kweon S-S, Kim M-G, Kang M-R, Shin M-H, Choi J-S. Difference of stage at cancer diagnosis by socioeconomic status for four target cancers of the National Cancer Screening Program in Korea: results from the Gwangju and Jeonnam cancer registries. *J Epidemiol* 2017; 27: 299–304.

24 Aizer AA, Wilhite TJ, Chen MH, Graham PL, Choueiri TK, Hoffman KE. Lack of reduction in racial disparities in cancer-specific mortality over a 20-year period. *Cancer* 2014; 120: 1532–9.

25 Vather R, Petrushnko W, Chapman D, Sammour T, Mor I, Warner R. Factors predictive of advanced stage of colorectal cancer at presentation - a bi-national study. *Colorectal Dis* 2020. https://doi.org/10.1111/codi.15137.

26 Henson KE, Elliss-Brookes L, Coupland VH et al. Data resource profile: National Cancer Registration Dataset in England. *Int J Epidemiol* 2020; 49: 16–16h.

27 Hirst Y, Stoffel S, Baio G, McGregor L, von Wagner C. Uptake of the english bowel (Colorectal) Cancer Screening Programme: an update 5 years after the full roll-out. *Eur J Cancer* 2018; 103: 267–73.

28 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; 61: 344–9.

29 Office for National Statistics. Available at https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion#ethnic-group. (accessed 9 June 2020).

30 Mork ME, You YN, Ying J et al. High prevalence of hereditary cancer syndromes in adolescents and young adults with colorectal cancer. *J Clin Oncol* 2015; 33: 3544–9.

31 Chen FW, Sundaram V, Chew TA, Ladabaum U. Advanced stage colorectal cancer in persons younger than 50 years not associated with longer duration of symptoms or time to diagnosis. *Clin Gastroenterol Hepatol* 2017; 15: 728–37.

32 Hardcastle JD, Chamberlain JO, Robinson MH et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet* 1996; 348: 1472–7.

33 Kronborg O, Fenger C, Olsen J, Jørgensen OD, Søndergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet* 1996; 348: 1467–71.

34 Brewster DH, Thomson CS, Hole DJ, Black RJ, Stroner PL, Gillis CR. Relation between socioeconomic status and tumour stage in patients with breast, colorectal, ovarian, and lung cancer: results from four national, population based studies. *BMJ* 2001; 322: 830–1.

35 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.

36 Moore KJ, Sussman DA, Koru-Sengul T. Age-specific risk factors for advanced stage colorectal cancer, 1981–2013. *Prev Chronic Dis*, 2018; 23: 1981–2013.

37 Andrew AS, Parker S, Anderson JC et al. Risk factors for diagnosis of colorectal cancer at a late stage: a Population-Based Study. *J Gen Intern Med* 2018; 33: 2100–5.

38 Saunders CL, Abel GA, El Turabi A, Ahmed F, Lyratzopoulos G. Accuracy of routinely recorded ethnic group information compared with self-reported ethnicity: evidence from the English Cancer Patient Experience survey. *BMJ Open* 2013; 3: e002882.

39 Evandrou M, Falkingham J, Feng Z, Vlachantoni A. Ethnic inequalities in limiting health and self-reported health in later life revisited. *J Epidemiol Community Health* 2016; 70: 653–62.

40 Bécares L. (2013). *Which ethnic groups have the poorest health? Ethnic health inequalities 1991 to 2011.* Manchester: Centre on Dynamics of Ethnicity (CoDE) Briefing, Manchester University.

41 Wang Z, Li L, GudaK CZ et al. Adverse clinical outcome associated with mutations that typify African American Colorectal Cancers. *J Natl Cancer Inst* 2016; 108: djw164.

42 National Cancer Intelligence Network Data Briefing. *Interpreting geographic variation in cancer stage.* Public Health England, 2014.

43 Richards M. (2018). *Report of the Independent Review of Adult Screening Programmes in England.* NHS England.

44 McGregor LM, Bonello B, Kerrison RS et al. Uptake of bowel scope (Flexible Sigmoidoscopy) screening in the English National Programme: the first 14 months. *J Med Screen* 2016; 23: 77–82.

45 Moss S, Mathews C, Day TJ et al. Increased uptake and improved outcomes of bowel cancer screening with a faecal immunochemical test: results from a pilot study within the national screening programme in England. *Gut* 2017; 66: 1631–44.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. STROBE checklist.

Table S2. Risk of advanced stage presentation, defined as Stage I/II versus Stage III/IV, by ethnic group and stratified by Index of Multiple Deprivation (IMD) quintile in the young-onset colorectal cancer cohort. [a]Adjusted for age, gender, English region and tumour location. [b]The chi-square test for heterogeneity of odds ratios was used to assess for interaction in the unadjusted analysis of ethnicity stratified by IMD quintile for each ethnic group. The likelihood ratio test was used to assess interaction in the adjusted analysis by comparing logistic regression models with and without an interaction term for ethnicity–IMD quintile and thus, only one P-value has been presented.

Table S3. Logistic regression model of demographic factors associated with advanced stage presentation, defined as Stages I/II versus Stages III/IV, in the young-onset colorectal cancer cohort after multiple imputation of stage.

Table S4. Logistic regression model of demographic factors associated with advanced stage presentation, defined as Stages I/II/III versus Stage IV, in the young-onset colorectal cancer cohort.

Table S5. Logistic regression model of demographic factors associated with advanced stage presentation, defined as Stage I versus Stages II/III/IV, in the young-onset colorectal cancer cohort.

Table S6. Risk of advanced stage presentation, defined as Stages I/II versus Stages III/IV, by ethnic group and stratified by Index of Multiple Deprivation (IMD) quintile in the older-onset colorectal cancer cohort. [a]Adjusted for age, gender, English region and tumour location. [b]The chi-square test for heterogeneity of odds ratios was used to assess for interaction in the unadjusted analysis of ethnicity stratified by IMD quintile for each ethnic group. The likelihood ratio test was used to assess interaction in the adjusted analysis by comparing logistic regression models with and without an interaction term for ethnicity–IMD quintile and thus, only one *P*-value has been presented.

Table S7. Logistic regression model of demographic factors associated with advanced stage presentation, defined as Stages I/II versus Stages III/IV, in the older-onset colorectal cancer cohort after multiple imputation of stage.

Table S8. Logistic regression model of demographic factors associated with advanced stage presentation, defined as Stages I/II/III versus Stage IV, in the older-onset colorectal cancer cohort.

Table S9. Logistic regression model of demographic factors associated with advanced stage presentation, defined as Stage I versus Stages II/III/IV, in the older-onset colorectal cancer cohort.