
Mr. A C Chambers BMBS MSc PhD MRCSEd 1, 2
Dr. S W Dixon MB BS MRCP 1
Dr. P White PhD 3
Prof. A C Williams PhD 1
Mr. M G Thomas BSc MB BS MS FRCS FRCS (gen) 2
Mr. D E Messenger BMedSc MBChB MSc FRCS 2

Affiliations
1. University of Bristol, School of Cellular and Molecular Medicine, University Walk, Bristol. BS8 1TD
2. Department of Colorectal Surgery, University Hospitals Bristol, Bristol Royal Infirmary, Upper Maudlin Street, Bristol. BS2 8HW.
3. University of West of England, Department of Engineering Design and Mathematics, Coldharbour Lane, Stoke Gifford, Bristol BS16 1QY

Corresponding Author: David Messenger, Department of Colorectal Surgery, University Hospitals Bristol, Bristol Royal Infirmary, Upper Maudlin Street, Bristol. BS2 8HW, David.Messenger@UHBristol.nhs.uk. Tel: 0117 342 2808.

Disclaimer: No conflicts of interest

Authorship credit: 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.

Funding: ACC - Medical Research Council Clinical Research Training Fellowship (MR/N001494/1), David Telling Trust Small Research Grant; Elizabeth Blackwell Institute Clinical Primer Grant. ACW - MRC Research Grant (MR/R017247/1). Funders were not involved in the conception, design, data analysis, writing or decision to submit for publication.

Abstract

Background
Evidence is emerging that colorectal cancer (CRC) incidence is increasing in young adults, but the descriptive epidemiology required to better understand these trends is currently lacking.

Method
A population-based cohort study was carried out of all adults aged 20-49 years diagnosed with CRC in England between 1974 and 2015. Data were extracted from the NCRAS database using ICD9/10 codes for CRC. Temporal trends in age-specific incidence rates (IRs) according to gender, anatomical subsite, index of multiple deprivation (IMD) quintile and geographical region were analysed using Joinpoint regression.

Results
A total of 56,134 new diagnoses of CRC were analysed. The most sustained increase in IR was in the 20-29 age group which is mainly driven by a rise in distal tumours. The magnitude of IR increases was similar in both genders and across Index of Multiple Deprivation quintiles, although the most pronounced increases in incidence were in the southern regions of England.

Conclusion
CRC should no longer be considered a disease of older people: changes in incidence rates should be used to inform future screening policy, preventative strategies and research agendas, as well as increasing public understanding that younger people need to be aware of the symptoms of CRC.
**Introduction**

CRC is a major cause of cancer related mortality and is the third most common cause of cancer death in the UK.\(^1\)\(^2\) Advances in the surgical and oncological management of CRC are the most likely explanation for the UK age-standardised mortality rate reducing from 49 to 27 per 100 000 person-years over the past 40 years.\(^2\)

Despite age-standardised incidence rates remaining static in the UK, as well as in other high human development index (HDI) nations,\(^3\) there is increasing evidence that incidence rates are increasing in adults under 50 years of age. A US study, using SEER data, revealed a doubling in the incidence rate of both colon and rectal cancers in patients aged between 20 and 54 years since 1974.\(^4\) Similar findings have been demonstrated in cohorts from Canada,\(^5\)\(^6\) Australia,\(^7\) New Zealand,\(^8\) and most recently Europe,\(^9\) suggesting that the underlying risk of CRC is increasing in young people.

While males are well recognised to have a higher incidence of colon and rectal cancer in older age groups, there is little difference in the incidence rates between men and women in adults under 40 years of age.\(^10\)\(^11\) UK data have shown that males have a higher proportion of rectal tumours, but that females have a higher proportion of right-sided tumours.\(^12\) However, data on anatomical subsite has not been linked to age-specific incidence trends in the UK population. Data from North America suggest that incidence rate increases have been driven by an increase in distal tumours,\(^4\)\(^6\) whereas European data suggest that incidence rate increases are more pronounced for colon cancer.

Socioeconomic status (SES) is associated with several important CRC risk factors.\(^13\)\(^15\) In the UK, data from Northern Ireland have shown no difference in age-standardised incidence between deprivation deciles,\(^16\) unlike in Scotland, where men from more deprived areas have been shown to have an increased incidence of CRC with evidence of an increasing deprivation gap over time.\(^17\)\(^18\) Previous studies have focused on SES as a risk factor for CRC incidence, but this has never been analysed in the context of recent changes in age-specific incidence trends in young adults.

Significant variations in the burden of disease exist between the nine regions of England, including variation in the age-standardised rate of years of life lost to CRC.\(^19\)\(^20\) Understanding if there is a socioeconomic and regional variation in incidence rate trends in the young population could help elucidate potential aetiological factors.

While data from the UK has been incorporated in recent Europe-wide population-based studies,\(^9\) a more detailed description of the epidemiology underlying the recent increase in CRC incidence in young adults trends is required. This is vital, as young adults typically present with more advanced tumours that carry a poorer prognosis and a more thorough knowledge of the descriptive epidemiology would help inform future preventative strategies. Therefore, the aim of this study was to determine temporal trends in incidence of colorectal cancer stratified by gender, anatomical subsite in the colorectum, socioeconomic status and geographical region of England.
Methods

Data sources
This study is reported according to the STROBE guidelines for epidemiological studies. Data were obtained on all patients diagnosed with CRC aged 20 years and above from 1974 to 2015 using data from the National Cancer Registration and Analysis Service (NCRAS) (Request ID: ODR1718_067). NCRAS is a UK-wide partnership operated by Public Health England (PHE) to collect data on all types of cancer, including CRC, occurring in the English population.

Procedures

ICD codes were used to identify all diagnoses of CRC. ICD 9 codes [colon 153.0-153.9 (excluding 153.5 - appendix tumour) and rectum 154.0 and 154.1] for CRC were used for diagnoses made between 1974 to 1994. ICD 10 codes [colon C18.0-C18.9 (excluding C18.1 – appendix tumour) and rectum C19 (recto-sigmoid) or C20 (rectum)] were used for diagnoses made between 1995 to 2015 (appendiceal adenocarcinomas were excluded and analysed separately - supplemental figure 1). For the purposes of this study, young adults were defined as those aged 20-49 years with cases grouped into three age groups based on age at diagnosis: 20-29 years, 30-39 years and 40-49 years.

Mid-year population estimates (MYPE) were obtained from the Office for National Statistics (ONS) to provide population data stratified by age. MYPEs in conjunction with the number of new diagnoses were used to calculate age-specific incidence density rates per 100 000 person-years, referred to hereafter as the age-specific incidence rate, for each age group using the formula given below.

\[
\text{Age-specific incidence rate} = \frac{\text{Number of new cases in age group}}{\text{Mid-year population estimate of age-group}}
\]

The European Standard Population 2013 (ESP 2013) was then used to derive age-standardised incidence rates for colon and rectal cancer for the overall dataset (20-49 years), in accordance with the methodology for direct-standardisation by the ONS.\(^{21}\)

\[
\text{Age-standardised incidence rate} = \frac{\sum (\text{ESP of age-group x age-specific rate})}{\sum \text{ESP of age-group}}
\]

CRC cases were further stratified by gender (using gender-specific population estimates from the ONS as above), anatomical subsite: either proximal (caecum to descending colon) and distal (sigmoid to rectum), geographical region (using region-based population estimates from the ONS from 1981 onwards) and Index of Multiple Deprivation (IMD) quintile (from 2001 onwards). IMD is an area-based metric that combines weighted information from seven domains: Income (weighting 22.5%), Employment (22.5%), Education (13.5%), Health (13.5%), Crime (9.3), Barriers to housing & services (9.3%) and Living environment (9.3%). Lower-layer Super Output Areas (LSOA; 32 844 in England) are given a value based on these domains. IMD
quintiles were calculated by ranking all LSOA from most to least deprived and then splitting this ranking into five equal groups (each quintile has 20% of the ranked areas).

**Statistical analysis**

Data analyses were performed using Joinpoint Regression Program 22 (National Cancer Institute (NCI), https://surveillance.cancer.gov/joinpoint/, version 4.7.0.0) to analyse the magnitude and direction of temporal trends in age-specific incidence rates according to gender, anatomical site, IMD quintile and geographical region. Permutation analysis of the log transformed incidence rates was used to fit a series of joined lines with a minimum of 0 and a maximum of 5 join points. A series of comparisons among fitted models ranging from 0 to 5 join points was then undertaken to select the best fit model. This procedure allowed estimation of the annual percentage change (APC) in incidence. The squared correlation coefficient ($R^2$) was used to estimate the goodness-of-fit of the Joinpoint regression models to provide an indication of the extent of agreement between modelled and observed values. Inspection of residuals under the models presented herein did not give cause for concern, i.e. standard errors appeared homoscedastic, free from serial correlation and without any unduly influential observations.

Age-period-cohort modelling (National Cancer Institute’s Age Period Cohort web tool, https://analysisistools.nci.nih.gov/apc) was used to assess the independent effects of age, period and cohort on CRC incidence rates.23 This was performed for all adults aged above 20 years. Data were inputted using three ten-year age groups (20-29, 30-39 and 40-49 for the Joinpoint regression modelling while four ten-year period groups (1976-1985, 1986-1995, 1996-2005, 2006-2015) were used for the age-period-cohort modelling as it was necessary to have age and time-period groups covering an equal timespan. Therefore, there were 11 birth cohorts starting in 1886 through to 1986 in ten-year bands. Reference values for the age-period-cohort model were arbitrarily chosen from the first cohort analysed (1976-1985). Data presented from this model were shown as incidence rate ratios (IRR) and 95% confidence intervals (CI) to assess cohort effects. Local drift was estimated by presenting age-specific net annual percentage change in incidence rates.
Results

Of the 1,145,639 new cases of CRC diagnosed between 1974 to 2015 in adults aged over 20 years, there were 2,594 cases in 20-29 year olds, 11,406 cases in 30-39 year olds, and 42,134 in 40-49 year olds.

Age-specific trends according to gender

Following an initial reduction in CRC incidence rates, there was a marked increase in rates among both 20-29 and 30-39 year olds. In 20-29 year-olds (figure 1A), incidence rate increases commenced earlier in females (APC=4.6% (95%CI 3.3 to 5.9%) from 1986) than in males (APC=5.1% (95%CI 3.7 to 6.5%) from 1992). In 30-39 year-olds (figure 1B), incidence rate increases commenced a decade later than in 20-29 year-olds with incidence rate increases again being observed earlier in females (APC=3.8% (95%CI 2.9 to 4.8%) from 1995) than in males (APC=6.0% (95%CI 4.4 to 7.6%) from 2002). The incidence rate trends observed in the younger age groups were more attenuated in 40-49 year olds (figure 1C), with small increases observed from 2003 onwards in both women (APC=1.5% (95%CI 0.5 to 2.5%)) and men (APC=0.8% (95%CI -0.1 to 1.6%).) These findings were suggestive of an age-cohort effect and assessed in more detail using age-period-cohort modelling applied to the entire adult population aged over 20 years. Using the 1926 birth cohort as the reference group, the incidence rate ratio (IRR) of CRC for cohorts born from 1886 to 1966 remained constant, following which there was a progressive increase in IRRs for successive birth cohorts (1976 cohort IRR=1.4, 95%CI 1.1 to 1.8; 1986 cohort IRR=2.2, 95%CI 1.3 to 3.8) (supplementary figure 1B-K).

Age-specific trends according to anatomical subsite

Increases in proximal cancer incidence rates were noted in 20-29 year olds (APC=4.4% (95%CI 2.3 to 6.5%) from 1995) and 30-39 year olds (APC=5.8% (95%CI 3.3 to 8.3%) from 2005), but with no observed effect in 40-49 year olds (APC=0.0% (95%CI -1.1 to 1.1%) from 2004) (figures 2A-C). The increase in proximal cancer age-standardised incidence rates among 20-49 year olds was predominantly driven by increases in the incidence of caecal and ascending colon cancers (supplemental figure 2). Age-specific incidence rate increases in distal cancers were more sustained and of a greater magnitude in comparison to proximal cancers among 20-29 year olds (APC=5.6% (95%CI 4.4 to 6.8%) from 1991) and 30-39 year olds (APC=3.3% (95%CI 1.0 to 5.7%) from 1995-2005 and APC=7.0% (95%CI 4.2 to 9.8%) from 2006). A less pronounced increase in distal cancer was also noted among 40-49 year olds (APC=1.4% (95%CI 0.7 to 2.1%) from 2001).

Age-standardised trends according to IMD quintile

The age-standardised incidence rates of distal cancers increased more rapidly than proximal cancers in all quintiles, except quintile 2 (supplemental figure 3A-E). There was no statistically significant difference in the magnitude of incidence rate increases across the quintiles for either proximal (p=0.110) or distal cancers (p=0.230).

Age-standardised trends according to geographical region

In 1985, age-standardised incidence rates of proximal cancers among 20-49 year olds were decreasing across all regions of England, except in London, with the greatest reduction observed in the South West (APC=-12.1%, 95%CI -20.3 to -3.1%) (figure 3).
By 2015, incidence rates were increasing the fastest in the south-eastern regions (APC South East=7.4%, 95%CI 4.8 to 10.1%; London=6.5%, 95%CI 0.1 to 13.2%; East of England=6.0%, 95%CI 2.5 to 9.7%). A similar, but more pronounced trend, was noted for distal cancers (figure 4). By 2005, the most rapid increase in distal cancer age-standardised incidence rates was noted in the South West (APC=10.1% (95%CI 6.1 to 14.1%) with all other southern regions experiencing annual increases of greater than 5%.

Discussion

This is the largest study based on a single, national population registry to describe detailed epidemiological changes in CRC incidence in a young adult population. The finding that CRC incidence is increasing rapidly in young adults supports recent findings from other high HDI nations. Rapid increases were observed in adults aged 20-39 years, which appears to be driven by increases in the rate of distal tumours. Incidence rate increases in the English population appear to be similar in both genders and across all socioeconomic groups. Importantly, incidence rates are increasing the fastest in the southern regions of England, particularly in the South West where the incidence of distal cancers is now increasing by more than 10% each year. A substantial birth cohort effect is observed with dramatic increases in IRRs from the mid-1960s onwards, similar to the observations in North American studies, although incidence rate ratio increases in these studies appear to have occurred in birth cohorts born approximately 15 years earlier. This suggests that any exposure to underlying risk factors may have occurred earlier in the North American population. Tumours in young adults are thought to be sporadic in nature, with environmental factors likely playing a significant causative role. The rising incidence of CRC in young adults coincides with several environmental changes most notably increasing childhood and adult obesity rates. It is recognised that early-life obesity leads to an increased risk of developing CRC. Therefore the increases in CRC incidence in young men and women may reflect the recent UK obesity prevalence trends, where prevalence rates among adults aged 35-54 years have increased from 15.4% to 26.3% in men and 17.9% to 24.5% in women, between 1993 and 2004. The more pronounced increase in the incidence rate of distal tumours compared to proximal tumours contrasts findings from recent European data, but is similar to the results from several North American studies. While risk factors associated with an increased risk of CRC have been identified, the strength of their association with tumour development at individual sites within the colorectum remains unclear. Differences in the way environmental factors promote tumorigenesis at various sites within the colorectum suggest that proximal and distal tumours may be biologically distinct entities; this may explain why the incidence in distal tumours from this English cohort has increased more rapidly. The biological differences in early versus late onset CRC have been explored by several studies: a recent large cohort study characterising the clinical and molecular features of early-onset CRC demonstrated enrichment of certain phenotypes such as consensus molecular subtype 1 (CMS1) in distal tumours in adults under 50 years. Other work has shown low levels of microsatellite
instability (MSI) in CRC in young adults. Additionally, there is a prevalence of mutations in genes such as β-catenin and KRAS. Interestingly, the combination of altered environmental exposures combined with the different tumour biology suggests that young adult CRC may be a different disease to later onset disease. This study showed no evidence for an association between SES and the rate of increase in incidence of both proximal and distal tumours, contrary to previous studies where higher incidence rates were observed in more deprived groups. Although factors associated with an increased risk of CRC, such as obesity, low fibre diet and reduced physical activity, are known to be associated with lower SES, changes in obesity prevalence trends are actually similar between socioeconomic groups and may partly explain the lack of association between SES and CRC incidence rate increases observed in this study. Additionally, obesity is one of many risk factors associated with the development of CRC and is itself caused by several complex societal, genetic and environmental interactions. It is perhaps not surprising that understanding the causative effects of single environmental risk factors is challenging.

Geographical inequalities in health are well characterised in England with incidence rates of all cancers noted to be higher in the North of England than in the South, although there is minimal variation in colorectal cancer incidence by region. In this study we observed recent incidence rate increases in CRC across all English regions, although the most marked increases were observed in the South of England. It is difficult to explain why incidence rates are increasing more rapidly in young adults in the South given that risk factors such as obesity are increasing faster in Northern regions. It is important to point out that the effect of regional variations in access to healthcare/endoscopy services on CRC incidence rates remains unknown and it may be that the observed incidence rate increases seen in the more affluent, southern regions are driven by increased awareness and access to medical care.

The main strengths of this study are the size and completeness of the dataset. Data were obtained from NCRAS, a nationally curated cancer registry, with 100% complete data for 1974-2012 and 98.4% complete data for 2013-2015. Unfortunately, stage-specific data were not routinely recorded until 2012, so further analysis of incidence rate trends according to tumour stage could not be performed. It will be important to know whether the increase in young-onset CRC was driven by an increase in the detection of early stage disease, particularly in regions and socioeconomic groups that may have increased health awareness and access to endoscopy services. Data presented in this study are population-based in nature and specific causal inferences cannot be made. In addition, IMD quintile and geographical region are group-level metrics and are unable to account for individual level contextual effects that could have affected the association between these variables and observed CRC incidence rates. Finally, with the increasing use of endoscopy in England, it could be argued that this accounted for the rising incidence of CRC. However, detection bias is unlikely as incidence rates were decreasing until the 1990s, and the most rapid increases were observed in the youngest age groups (the least likely to attend for endoscopic examination).
In summary, the incidence rate of young-onset CRC cancer is increasing, particularly among adults aged 20-39 years. This trend appears to be predominantly driven by a rise in distal tumours. Incidence rate increases of a similar magnitude have been observed in both genders and across IMD quintiles, but are most pronounced in the South of England. Importantly, there is a strong birth cohort effect and it is likely that the increased risk in the youngest cohorts will be carried forward as they age, which will place a significant burden on future healthcare resources. The role of environmental factors such as diet, obesity, physical exercise and the gut microbiota in the development of young-onset CRC are incompletely understood and require further research. Reducing the screening age below 50 years will have significant resource implications in the current economic climate and instead, there should be more focus on risk stratifying symptomatic younger patients to further investigation using tests such as quantitative faecal immunohistochemical testing.

Role of the funding source
Funders played no role in study design; in the collection, analysis and interpretation of these data; in the writing of the report; or in the decision to submit the paper for publication

Declaration of Competing Interests
No financial or personal declarations of competing interest

Contributors statement
ACC literature search, study design, data analysis, writing, figures, data interpretation; SWD- data analysis, data interpretation, writing; PW- data analysis, data interpretation, writing; ACW- data interpretation, writing; MGT- data interpretation, writing; DEM- literature search, study design, data interpretation, writing.

Acknowledgements
ACC - Medical Research Council Clinical Research Training Fellowship (MR/N001494/1), David Telling Trust Small Research Grant; Elizabeth Blackwell Institute Clinical Primer Grant. ACW - MRC Research Grant (MR/R017247/1).
References


