Measurement of clinical delay intervals among younger adults with colorectal cancer using health administrative data: a population-based analysis

Matthew Castelo 1,2,3, Adena S Scheer 1,2,4, Lawrence Paszat 2,3, Bettina E Hansen 2, Neil Faught 3, Lena Nguyen 3, Nancy N Baxter 1,2,3,4,5

ABSTRACT

Background Clinical delays may be important contributors to outcomes among younger adults (<50 years) with colorectal cancer (CRC). We aimed to describe delay intervals for younger adults with CRC using health administrative data to understand drivers of delay in this population.

Methods This was a population-based study of adults <50 diagnosed with CRC in Ontario, Canada from 2003 to 2018. Using administrative code-based algorithms (including billing codes), we identified four time points along the pathway to treatment—first presentation with a CRC-related symptom, first investigation, diagnosis date and treatment start. Intervals between these time points were calculated. Multivariable quantile regression was performed to explore associations between patient and disease factors with the median length of each interval.

Results 6853 patients aged 15–49 were diagnosed with CRC and met the inclusion criteria. Males comprised 52% of the cohort, the median age was 45 years (IQR 40–47), and 25% had stage IV disease. The median time from presentation to treatment start (overall interval) was 109 days (IQR 55–218). Time between presentation and first investigation was short (median 5 days), as was time between diagnosis and treatment start (median 23 days). The greatest component of delay occurred between first investigation and diagnosis (median 78 days).

Conclusions Some younger CRC patients experience prolonged times from presentation to treatment, and time between first investigation to diagnosis was an important contributor. Access to endoscopy may be a target for intervention.

INTRODUCTION

The incidence of colorectal cancer (CRC) is increasing in younger adults (<50 years)—this trend has been identified in Canada, the United States, Italy, Australia and other nations. 1–4 CRC mortality in this population is also rising, and these patients are more likely to be diagnosed with advanced and metastatic disease compared with older adults. 5–7 Only 16% of younger adults will have an identifiable predisposing factor and therefore most will not be eligible for screening. 8

Delays to diagnosis and treatment have been suggested as possible contributors to poor outcomes. 8–10 However, the literature reporting delay intervals in younger CRC patients is heterogeneous and of variable quality. The Aarhus statement is an international consensus document that has been published in an effort to improve the methodology of cancer delay studies and includes standardised definitions of delay intervals. 11

There is a paucity of large, population-based studies, and few examine pre-diagnostic intervals, which may be particularly important in
a younger population. Pre-diagnostic intervals include time occurring between the development of cancer-related symptoms and diagnosis. However, it is challenging to identify pre-diagnostic time points such as the date of first presentation using health administrative data. Important advances in methodology have been made, including algorithms developed for population-based databases. These newer approaches have not focused specifically on younger adults with CRC, and important intervals in this group remain understudied. Time from the date of first presentation with CRC-related symptoms to the date of first investigation (the doctor interval) may be a source of delay for younger CRC patients if they experience misdiagnosis or do not receive timely access to initial investigations. Previous studies have not reported this interval for younger CRC patients and have been limited in their ability to understand the relative contributions of individual intervals to the overall delay experienced.

The aim of this study was to describe important time points and delay intervals along the pathway to treatment for adults <50 with CRC using population-based health administrative data and associated factors to understand potential drivers of longer intervals in this population.

METHODS
Study design and data sources
This was a population-based cohort study of adults <50 years diagnosed with CRC in Ontario, Canada. Data were obtained from ICES (formerly known as the Institute for Clinical Evaluative Sciences), an independent, non-profit research institute that maintains health administrative data for more than 14 million Ontario residents and provides de-identified linked datasets to researchers. Using a unique encrypted identifier, patient data were linked from 15 separate data sources and analysed at ICES (online supplemental material 1). We followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement and the Aarhus checklist, a quality assessment tool for cancer delay studies in the preparation of this manuscript (online supplemental material 2). The Research Ethics Board at the University of Toronto (number 41634) and St. Michael’s Hospital (number 21-207) approved this study.

Patient population
Using the Ontario Cancer Registry (OCR), we identified Ontario residents aged 15–49 with valid Ontario Health Insurance Plan (OHIP) numbers who were diagnosed with CRC (online supplemental material 3) between 1 October 2003 and 31 December 2018. This time frame accounted for a change in administrative data in Ontario from International Classification of Diseases (ICD)–9 to ICD-10 beginning in 2001. The OCR is a high-quality population-based cancer registry operated by Cancer Care Ontario since 1964, with over 94% of included CRCs being microscopically confirmed. Exclusion criteria included previous malignancy in the OCR, CRC diagnosis on death certificate only, appendiceal cancer and inflammatory bowel disease (IBD). Patients with IBD were identified and excluded as we anticipated this population to have a unique diagnostic journey with some eligible for CRC surveillance (online supplemental material 3).

Outcomes
Using detailed health administrative data, we identified four distinct time points along the pathway to treatment—first presentation with a CRC-related symptom, first investigation, diagnosis date and treatment start. This enabled us to measure five overlapping delay intervals (figure 1). These included the doctor interval (presentation to first investigation), system interval (first investigation to treatment start), diagnostic interval (presentation to diagnosis) and treatment interval (diagnosis to treatment start) as defined by the Aarhus statement. We also calculated the overall interval (presentation to treatment start).

While not intervals recognised in the Aarhus statement, we additionally determined the time from presentation to the first abdominal/pelvic CT or MRI, and the time from presentation to first lower endoscopy for descriptive purposes.

Assigning date of first presentation
In administrative data, to accurately identify the date of first presentation for CRC-related symptoms, which may be non-specific, these visits must be distinguished from routine encounters with the healthcare system. We
assigned the date of first presentation with CRC-related signs and symptoms using an algorithm adapted from prior work by Groome et al.,\textsuperscript{15, 23} who defined date of first presentation based on procedural and diagnostic codes that occurred more frequently prior to diagnosis compared with a background period. Codes were organised into categories, and category-specific lookback periods were established based on when the frequency of codes was similar to the background rate (online supplemental material 4). This method has been used for oral cancer, breast cancer and CRC\textsuperscript{15, 23 24} and is being tested for use in pancreatic cancer.\textsuperscript{25} The methodology avoids arbitrary lookback periods, leverages existing health administrative datasets and allows for a more complete assessment of the diagnostic journey.\textsuperscript{25}

Based on this algorithm, we assigned the date of first presentation for each patient using the following steps:

1. For each patient in our cohort, all administrative codes occurring in the OHIP, National Ambulatory Care Reporting System, Same-Day Surgery and Discharge Abstract Database databases up to 18 months prior to diagnosis were identified.
2. Codes included in the algorithm were retained and organised by date.
3. Codes were assigned into their respective categories, and any code recorded at a time exceeding the category-specific lookback period was removed. The date of the earliest code was identified (online supplemental material 4).
4. Codes included in categories 1–6 were ICD-10 and OHIP diagnostic codes (online supplemental file 4) and do not contain information on referring physicians. If the earliest code was included in one of these categories, this was the date of first presentation.
5. Code categories 7–19 (online supplemental material 4) contained procedural codes and OHIP billing codes for endoscopy, endoscopic procedures, surgery, specialist consultations and imaging. For these codes, referring physician information was available. If any codes in these categories were identified as the earliest code, the referring physician was identified, and the most recent visit with that physician (within 18 months of diagnosis) was the date of first presentation. If the referring physician was missing, the most recent encounter with the performing or ordering physician was used (within 18 months of diagnosis). If both the referring and performing physician were missing, the code date itself was the date of first presentation.

We performed sensitivity analyses around the date of first presentation. We explored four alternative algorithms to assign date of presentation (online supplemental material 5). We used the original set of administrative codes but removed the category-specific lookback dates from the CRC diagnosis date, instead testing a universal 12-month cut-off and a universal 18 month cut-off. We examined frequency tables of codes occurring in our cohort of younger CRC patients, and identified an additional 201 administrative codes plausibly related to CRC that were not in the original algorithm (online supplemental material 4). This expanded algorithm was tested with a 12 and 18 month cut-off from diagnosis date.

Assigning remaining time points
The date of first investigation was defined as the first occurrence of an intervention or diagnostic test after the date of first presentation (online supplemental material 4). These included endoscopy, faecal occult blood test, imaging and blood work. Codes used to determine the date of presentation were also used for the date of first investigation when these codes represented the above tests or interventions. The date of diagnosis was determined from the OCR, which follows a standardised hierarchical method of assigning cancer diagnosis dates; the earliest among admission date (for hospitalisation), registration date (for emergency visits), diagnosis date in the cancer centre records or collection date for pathology confirmation. Finally, the date of first treatment was defined as the first date of any chemotherapy administration, radiotherapy treatment or surgical intervention after diagnosis (online supplemental material 4).

Covariates
Age at diagnosis and sex were determined from the Registered Persons Database.\textsuperscript{26} Marginalisation was measured using the Ontario Marginalization Index (ON-Marg), which is a validated tool that measures residential instability, material deprivation, dependency, and ethnic concentration at the neighbourhood level.\textsuperscript{27} The ON-Marg contains a number of socioeconomic variables in its derivation, including average number of persons per dwelling, residential mobility, per cent lone-parent families, education, unemployment, per cent recent immigrants, among others.\textsuperscript{27} A summary score was created by averaging each individual’s quintile for the four measures, with higher scores indicating greater levels of marginalisation.\textsuperscript{27} Canadian census data were used to assign individuals into income quintiles (1 being the least affluent and 5 being the most affluent), based on their postal code of residence.\textsuperscript{28} Rural patients were considered in a separate category of income, defined as those living in communities ≤10,000 individuals in size. We did not include both ON-Marg and income quintile in the same model as they are derived using income data from the same source. Disease characteristics were obtained from the OCR, including stage, cancer site, and cause of death. Stage was available for patients diagnosed after 2007. Access to care was assessed by calculating the use of lower gastrointestinal (GI) endoscopy (colonoscopy, sigmoidoscopy or anoscopy) over the year preceding CRC diagnosis in the patient’s Local Health Integration Network (LHIN), a geographic unit of health administration in Ontario.

Statistical analysis
Patient characteristics were described. Differences between age groups (15–39, 40–44 and 45–49 years) were
RESULTS

Patient and disease characteristics
From the OCR, 8081 patients aged 15–49 were diagnosed with CRC between 2003 and 2018. After exclusions, 7053 patients remained. We were unable to assign a date of first presentation for 200 patients, resulting in 6853 patients included in the analysis (online supplemental material 6). Patient characteristics are presented in table 1. Men comprised 52% (3,587) of the cohort, and the median age was 45 years (IQR 40–47). Rectal cancer patients represented 31% of the cohort. For those diagnosed after 2007 when stage of disease was available, 25% of patients presented with stage IV disease. Patients aged 15–39 were more likely to have stage IV disease (28% age 15–39, 24% age 40–44, 25% age 45–49; p<0.001) and proximal colon cancer (39% age 15–39, 32% age 40–44, 34% age 45–49; p<0.001) compared with older patients (table 1).

Patient characteristics stratified by sex are presented in online supplemental material 7.

Characterising dates of first presentation, investigation and diagnosis
Codes defining the date of first presentation, investigation and diagnosis were visualised and explored (figure 2). The date of first presentation was characterised mainly by physician encounters for abdominal pain, diarrhoea, constipation/obstruction, and rectal bleeding. Additionally, codes for abdominal X-ray and general surgery consultations were common. The date of first investigation consisted of abdominal imaging, endoscopy and more general interventions such as blood work and chest X-rays. The diagnosis date was most often characterised by endoscopy and endoscopic biopsies.

Delay intervals
Descriptive statistics for the five intervals are presented in figure 3, table 1. The median time from presentation to treatment (overall interval) was 109 days (IQR 55–218). The overall interval can be divided into the diagnostic and treatment intervals. Most of this time was contained within the diagnostic interval (median 78 days, IQR 28–186). Median time from diagnosis to treatment (treatment interval) was 23 days (IQR 7–40). The overall interval can also be divided into the doctor and system intervals. In our cohort, we found the doctor interval was a median of 5 days (IQR 0–36) and the system interval was a median of 77 days (IQR 36–160). The alternative algorithms produced very similar intervals (median different of 0 days for all intervals).

There were weak correlations between contiguous intervals (doctor/system intervals and diagnostic/treatment intervals; online supplemental material 8). Longer doctor intervals were weakly and negatively correlated with system intervals (correlation coefficient −0.09; p<0.001). There was no significant correlation between diagnostic interval and treatment interval (correlation coefficient=0.02; p=0.085).

The median time from presentation to first abdominal/pelvic CT or MRI was 46 days (IQR 7–142), and median time from presentation to first lower endoscopy was 70 days (IQR 28–163). There was not a significant difference by age for these measures (table 1).

Factors associated with delay intervals
Table 2 presents adjusted quantile regression for the delay intervals of interest. Univariate regression is presented in online supplemental material 9 and 10. Older age was not significantly associated with the overall interval, doctor interval, or system interval. None of the relationships was found to be significantly non-linear (all p>0.05), and age was modelled as a simple continuous covariate (online supplemental material 11). Increasing age was associated with shorter diagnostic intervals (median difference −3.5 days per 5 year increase, 95% CI −6.4 to −0.6) and longer treatment intervals (median difference 2.0 days per 5 year increase, 95% CI 1.4 to 2.5). Compared with men, women had significantly longer overall intervals (median difference 26.5 days, 95% CI 18.9 to 34.2). More contemporary diagnosis was associated with a statistically significant increase in the treatment interval (median increase 0.5 days per year, 95% CI 0.2 to 0.8) and a statistically significant decrease in the doctor interval (median decrease 0.2 days per year, 95% CI −0.3 to 0). Patients with rectal cancer had longer overall intervals compared with those with proximal colon cancer (median difference 22.1 days, 95% CI 13.2 to 30.9), and most of this
| Characteristic                              | Overall | 15–39 years | 40–44 years | 45–49 years | P-value |
|--------------------------------------------|---------|-------------|-------------|-------------|---------|
| Age (years)                                | 45.0 (40.0, 47.0) | 35.0 (31.0, 38.0) | 42.0 (41.0, 43.0) | 47.0 (46.0, 49.0) | <0.001 |
| Sex                                        |         |             |             |             | 0.4    |
| Male                                       | 3587 (52%) | 809 (53%) | 931 (51%) | 1847 (53%) |        |
| Female                                     | 3266 (48%) | 731 (47%) | 896 (49%) | 1639 (47%) |        |
| ON-Marg Summary Score                      | 3.00 (2.25, 3.50) | 3.00 (2.50, 3.75) | 2.75 (2.25, 3.50) | 3.00 (2.25, 3.50) | 0.017  |
| Missing                                    | 76      | 13          | 15          | 48          |        |
| Income quintile                            |         |             |             |             | 0.009  |
| 1                                          | 1174 (17%) | 272 (18%) | 310 (17%) | 592 (17%) |        |
| 2                                          | 1154 (17%) | 290–300* (19%) | 290–300* (16%) | 568 (16%) |        |
| 3                                          | 1234 (18%) | 289 (19%) | 352 (19%) | 593 (17%) |        |
| 4                                          | 1298 (19%) | 289 (19%) | 353 (19%) | 656 (19%) |        |
| 5                                          | 1231 (18%) | 252 (16%) | 332 (18%) | 647 (19%) |        |
| Rural                                      | 748 (11%) | 139 (9.0%) | 190 (10%) | 419 (12%) |        |
| Missing                                    | 14      | <10*        | <10*        | 11          |        |
| Stage (percentages reflect diagnoses 2007 and later) |         |             |             |             | <0.001 |
| I                                          | 884 (17%) | 161 (13%) | 253 (18%) | 470 (18%) |        |
| II                                         | 1082 (21%) | 243 (20%) | 309 (22%) | 530 (20%) |        |
| III                                        | 1939 (37%) | 465 (39%) | 491 (35%) | 983 (37%) |        |
| IV                                         | 1321 (25%) | 337 (28%) | 338 (24%) | 646 (25%) |        |
| Pre-2007                                   | 1320    | 264         | 360         | 696         |        |
| Missing                                    | 307     | 70          | 76          | 161         |        |
| Disease site                               |         |             |             |             | <0.001 |
| Proximal colon                             | 2363 (34%) | 603 (39%) | 590 (32%) | 1170 (34%) |        |
| Rectum                                     | 2103 (31%) | 461 (30%) | 564 (31%) | 1078 (31%) |        |
| Sigmoid and rectosigmoid                   | 2387 (35%) | 476 (31%) | 673 (37%) | 1238 (36%) |        |
| Use of lower Gl endoscopy in patient LHIN (procedures per 100 000 residents in last year) | | | | | 0.029 |
| 4181 (3411, 4648)                          | 4208 (3482, 4681) | 4179 (3401, 4639) | 4166 (3369, 4632) |        |
| Time from presentation to first abdominal/pelvic CT or MRI (days) |         |             |             |             | 0.6    |
| Median (IQR)                               | 46 (7, 142) | 47 (7, 165) | 47 (7, 140) | 46 (7, 130) |        |
| Mean (SD)                                  | 92 (108) | 98 (114) | 92 (109) | 88 (105) |        |
| Range                                      | 0, 500 | 0, 486 | 0, 500 | 0, 482 |        |
| Missing                                    | 4221 | 866 | 1129 | 2226 |        |
| Time from presentation to first lower endoscopy (days) |         |             |             |             | 0.071  |
| Median (IQR)                               | 70 (28, 163) | 79 (27, 182) | 70 (30, 172) | 67 (28, 154) |        |
| Mean (SD)                                  | 111 (108) | 117 (112) | 113 (109) | 107 (106) |        |
| Range                                      | 0, 493 | 0, 486 | 0, 456 | 0, 493 |        |
| Missing                                    | 1680 | 451 | 416 | 813 |        |
| Doctor interval (presentation to first investigation; days) |         |             |             |             | 0.9    |
| Median (IQR)                               | 5 (0, 36) | 4 (0, 42) | 4 (0, 36) | 5 (0, 35) |        |
| Mean (SD)                                  | 36 (68) | 38 (70) | 36 (68) | 34 (66) |        |

Continued
difference was contained within the treatment interval (median difference 19.6 days, 95% CI 17.9 to 21.4). More frequent use of lower GI endoscopy in the patient’s LHIN was associated with significantly shorter overall, diagnostic and system intervals.

More advanced stage was consistently associated with shorter delay intervals. Compared with those with stage I disease, the median time from presentation to treatment for patients with stage IV disease was more than 2 months shorter (adjusted median difference for overall interval −61.3 days, 95% CI −72.5 to −50.2). Differences were also apparent between patients with stage I and stage II or stage III disease (stage II vs stage I overall interval median difference −34.3 days, 95% CI −48.1 to −20.4; Stage III versus stage I overall interval median difference −41.7 days, 95% CI −54.1 to −29.3).

**DISCUSSION**

In this population-based cohort study, 6853 adults <50 years diagnosed with CRC between 2003 and 2018 in Ontario were found to have median intervals from first presentation to first treatment (overall interval) of 109 days. More than 25% of these patients had overall intervals greater than 218 days (7.2 months). We found the greatest contributor to the overall interval was time from first investigation to treatment (median 77 days), while the median time from diagnosis to treatment start was 23 days. We found initial investigations occurred a median 5 days after presentation in younger CRC patients. Among patients with stage available, 25% presented with stage IV disease, and more advanced stage was strongly associated with shorter intervals. Factors associated with longer overall intervals included female sex and distal tumours.

This is the first study to identify the date of initial investigation using health administrative data and report the doctor interval and system interval among younger adults with CRC. We found the median doctor interval was short (5 days), suggesting when younger patients seek medical attention, the initial workup is begun in a timely fashion for most patients. However, most of these consisted of blood work, ultrasound, and X-rays. Time to more definitive investigations, including CT, MRI or endoscopy took longer. After diagnosis, the median time to treatment was also within recommended cutoffs of 1 month for most patients, although there was a small increase in this interval over the study period. However, our results demonstrate that the period of time most contributory to overall wait time was between the first investigation and

| Characteristic          | Overall N=6853 | 15–39 years N=1540 | 40–44 years N=1827 | 45–49 years N=3486 | P-value |
|-------------------------|---------------|-------------------|-------------------|--------------------|---------|
| Range                   | 0, 462        | 0, 456            | 0, 409            | 0, 462             |         |
| Missing                 | 99            | 31                | 18                | 50                 |         |
| System interval (first investigation to treatment; days) |               |                   |                   |                    | 0.2     |
| Median (IQR)            | 77 (36, 160)  | 78 (34, 168)      | 80 (39, 163)      | 74 (37, 155)       |         |
| Mean (SD)               | 114 (107)     | 116 (109)         | 117 (108)         | 112 (105)          |         |
| Range                   | 0, 763        | 0, 601            | 0, 638            | 0, 763             |         |
| Missing                 | 470           | 108               | 108               | 254                |         |
| Diagnostic interval (presentation to diagnosis; days) |               |                   |                   |                    | 0.002   |
| Median (IQR)            | 78 (28, 186)  | 89 (28, 208)      | 83 (30, 195)      | 72 (27, 176)       |         |
| Mean (SD)               | 121 (117)     | 129 (121)         | 124 (117)         | 115 (115)          |         |
| Range                   | 0, 519        | 0, 503            | 0, 519            | 0, 502             |         |
| Missing                 | 384           | 82                | 91                | 211                |         |
| Treatment interval (diagnosis to treatment; days) |               |                   |                   |                    | <0.001  |
| Median (IQR)            | 23 (7, 40)    | 20 (4, 35)        | 24 (8, 41)        | 25 (8, 42)         |         |
| Mean (SD)               | 29 (31)       | 25 (31)           | 29 (30)           | 30 (32)            |         |
| Range                   | 0, 357        | 0, 343            | 0, 357            | 0, 350             |         |
| Missing                 | 384           | 82                | 91                | 211                |         |
| Overall interval (presentation to treatment; days) |               |                   |                   |                    | 0.052   |
| Median (IQR)            | 109 (55, 218) | 113 (51, 233)     | 114 (58, 226)     | 106 (56, 208)      |         |
| Mean (SD)               | 148 (121)     | 152 (125)         | 153 (121)         | 145 (120)          |         |
| Range                   | 0, 763        | 0, 631            | 0, 638            | 0, 763             |         |
| Missing                 | 384           | 82                | 91                | 211                |         |

Presented are median (IQR) or N (%) unless otherwise specified. *Ranges are given so as to suppress small cells of missing data. GI, gastrointestinal; LHIN, local health integration network.
diagnosis. For most patients, the date of diagnosis corresponded to the date of endoscopy, indicating that access to endoscopy was a key driver of the overall wait time for patients.

The literature reporting delay intervals in younger adults with CRC is heterogeneous.\(^7\)\(^\text{15-18}^\text{38-39}\) Webber et al\(^\text{15}\) established the algorithm used in this study to identify the date of first presentation and examined the diagnostic interval among 23,961 adults (mean age 68.6 years) with CRC in Ontario. There were 1902 adults <50 years in the study, and the median diagnostic interval was 111.5 days in those <35 years, decreasing to 73 days in patients aged 45–49. An additional Canadian study by Sikdar et al\(^\text{39}\) examined 822 younger CRC patients in Alberta, demonstrating a median diagnostic interval of 81 days. Finally, Price et al\(^\text{38}\) studied 1206 younger adults with CRC in the UK using health administrative data and reported the diagnostic interval to be a median of 65 days. Despite using different population-based data sources and a variety of algorithms for determining the date of first presentation, these results are very similar to our result for the diagnostic interval (median 78 days).

Our findings regarding the treatment interval are also in keeping with previous large reports. In a very large analysis of the National Cancer Database (NCDB), Gabriel et al\(^\text{7}\) found the mean time from diagnosis to treatment was 11.2 days among 94,143 younger colon cancer patients, and 22.0 days among 58,947 younger rectal cancer patients. The mean treatment interval was 29 days for our cohort, and we similarly found rectal cancer patients had longer treatment intervals. We identified only two studies reporting the overall interval (presentation to treatment start), both including less than 100 younger CRC patients.\(^40\)\(^\text{41}\) This literature shows many younger patients will wait longer than 2 months from presentation to diagnosis with CRC and some will have much longer delays (>6 months) but will subsequently receive treatment in a timely fashion.

Patients who experienced longer intervals were more likely to have early stage disease. This is consistent with previous work. Webber et al\(^\text{15}\) similarly showed CRC patients with stage I disease had the longest intervals (median 104 days for the diagnostic interval) compared with those with higher stage disease (stage IV median 62 days for the diagnostic interval), although this included mainly older adults. Among 253 younger adults in the USA, Chen et al\(^\text{42}\) demonstrated CRC patients with early stage disease had longer times for three intervals (symptom onset to presentation, presentation to diagnosis and symptom onset to diagnosis). This phenomenon has been recognised in cancer populations more broadly than just those with CRC, where it is hypothesised tumours that are more aggressive within the same cancer type may be more likely to cause bothersome symptoms and lead patients to present more quickly, thus shortening the interval to diagnosis.\(^43\) This U-shaped relationship between interval length and disease severity/mortality has been demonstrated in patients with colorectal, lung, melanoma, breast and prostate cancer.\(^44\)

Strengths of this study include the population-based nature of the data and large sample size. Our cohort included a broad sample of younger patients with CRC,
including 25% with metastatic disease at presentation. We used previously published algorithms to identify the date of first presentation\textsuperscript{15 23 24} and extended this work to include the date of first investigation. This allowed us to report novel intervals among younger CRC patients that have not previously been explored (doctor interval and system interval). The algorithms used were able to assign dates of presentation and investigation in over 95% of patients, demonstrating their utility in future research.

Our analysis was guided by international consensus-derived guidelines for the definitions of delay intervals (Aarhus statement,\textsuperscript{11} and we used sensitivity analyses to explore the date of first presentation. Finally, we used quantile regression to examine factors associated with intervals, which is an appropriate nonparametric method for analysing skewed data such as wait times.\textsuperscript{15 30–34}

This study has limitations. While the algorithms used to identify the date of first presentation have been previously reported and published, they are not validated against patient charts directly.\textsuperscript{15 23 24} Thus, there is the possibility for misclassification. We explored the cohort’s healthcare encounters prior to CRC diagnosis and identified an additional 201 administrative codes that were not present in the original algorithm in a sensitivity analysis. We also

---

**Figure 3**  Length of delay intervals and time from presentation to first abdominal/pelvic CT or MRI, and from presentation to first lower endoscopy in a cohort of colorectal cancer patients <50 years.
Table 2  Multivariable quantile regression for delay intervals of interest with patient and disease characteristics

| Characteristic                        | Overall interval | Overall interval | System interval | Overall interval | Overall interval | Overall interval |
|--------------------------------------|------------------|------------------|-----------------|------------------|------------------|------------------|
|                                     | Median difference (N=4982) | P-value | Median difference (N=5113) | P-value | Median difference (N=4926) | P-value | Median difference (N=5174) | P-value | Median difference (N=4982) | P-value |
| Age (increase in 5 years)            | −3 (−6.5 to 0.5) | 0.098           | 0 (−0.3 to 0.4) | 0.969           | −0.7 (−3 to 1.6) | 0.536           | −3.5 (−6.4 to −0.6) | 0.017 | 2 (1.4 to 2.5) | <0.001 |
| Sex                                  |                  |                 |                 |                 |                  |                 |                  |                 |                 |         |
| Male                                 | Ref              |                 | Ref              |                 | Ref              |                 | Ref              |                 | Ref              |         |
| Female                               | 26.5 (18.9 to 34.2) | <0.001 | 0.7 (−0.2 to 1.7) | 0.138 | 20.2 (14.6 to 25.8) | <0.001 | 27.2 (20.2 to 34.3) | <0.001 | 1.6 (0.3 to 3) | 0.021  |
| ON−Marg Summary Score (increase in one point) | 3.7 (−0.7 to 8.1) | 0.096 | −0.3 (−0.7 to 0.2) | 0.21 | 3 (−0.7 to 6.7) | 0.115 | 5 (0.6 to 9.4) | 0.024 | −0.1 (−1.1 to 0.9) | 0.873  |
| Disease site                         |                  |                 |                 |                 |                  |                 |                  |                 |                 |         |
| Proximal colon                       | 14.3 (5.6 to 23) | 0.001 | 1.4 (0.3 to 2.5) | 0.016 | 8.5 (1.6 to 15.4) | 0.015 | 5.4 (−2.5 to 13.4) | 0.18 | 8.3 (6.5 to 10.2) | <0.001 |
| Sigmoid and rectosigmoid             |                  |                 |                 |                 |                  |                 |                  |                 |                 |         |
| Rectum                               | 22.1 (13.2 to 30.9) | <0.001 | 4.9 (3.6 to 6.3) | <0.001 | 13.1 (6.3 to 19.8) | <0.001 | 5.5 (−2.6 to 13.5) | 0.184 | 19.6 (17.9 to 21.4) | <0.001 |
| Stage                                |                  |                 |                 |                 |                  |                 |                  |                 |                 |         |
| I                                    | −34.3 (−48.1 to −20.4) | <0.001 | −7.3 (−9.1 to −5.4) | <0.001 | −30 (−39.4 to −20.5) | <0.001 | −26.1 (−38.6 to −13.5) | <0.001 | −11.1 (−13.9 to −8.2) | <0.001 |
| II                                   | −41.7 (−54.1 to −29.3) | <0.001 | −5.9 (−8 to −3.9) | <0.001 | −30.7 (−39.6 to −21.9) | <0.001 | −35.2 (−45.7 to −24.7) | <0.001 | −12.6 (−15.2 to −10) | <0.001 |
| III                                  | −61.3 (−72.5 to −50.2) | <0.001 | −7.8 (−9.3 to −6.3) | <0.001 | −45.7 (−54.6 to −36.8) | <0.001 | −54.6 (−65.7 to −43.5) | <0.001 | −15 (−17.6 to −12.5) | <0.001 |
| IV                                   |                  |                 |                 |                 |                  |                 |                  |                 |                 |         |
| Use of lower GI endoscopy in patient LHIN (increase in 1000 procedures per 100 000 residents per year) | −7.6 (−13.8 to −1.4) | 0.017 | 0.4 (−0.4 to 1.2) | 0.328 | −7.6 (−13.1 to −2.1) | 0.007 | −7.3 (−14.3 to −0.3) | 0.041 | −1 (−2.5 to 0.5) | 0.172  |
| Year of diagnosis (increase in 1 year) | 0.3 (−1 to 1.6) | 0.629 | −0.2 (−0.3 to 0) | 0.046 | 0.7 (−0.3 to 1.8) | 0.159 | −0.2 (−1.6 to 1.1) | 0.723 | 0.6 (0.3 to 0.9) | <0.001 |

Effect estimates are presented as median days with 95% CIs. Models are adjusted for age, sex, ON-Marg score, disease site, access to lower endoscopy, year of diagnosis, density of specialists and stage (diagnoses in 2007 and later). GI, gastrointestinal; LHIN, local health integration network.
varied the lookback periods within the algorithm, and these sensitivity analyses did not change the results of the study. Our results are also comparable to those reported by other high-quality population-based studies on delay in younger adults with CRC using different methods and data sources, providing some degree of external validation.² ¹⁵ ³⁸ ³⁹ While the period between the onset of symptoms and first presentation is highly important in determining the overall length of delay experienced, we were unable to ascertain this interval with health administrative data. Finally, the data used in this analysis do not contain the indication for diagnostic tests and thus investigations such as blood work may not have been conducted as part of investigation of CRC signs or symptoms.

This large population-based study has described five delay intervals among patients <50 years of age with CRC in Ontario, Canada. Some patients in this study experienced prolonged times to treatment, and an important contributor to delay is the time from first investigation to diagnosis. Given the diagnosis of CRC was commonly made with endoscopy, access to this service is an important target for future research, and potentially intervention for younger patients. Further, we have identified an apparent disparity in outcomes between men and women, with young women experiencing a median 26.5 days longer time to treatment from presentation warranting further research.

Contributors MC planned the study, completed the analysis, wrote the manuscript and submitted the study. LP, BEH and AS planned the study, reviewed the analysis and reviewed the manuscript. NF and LN prepared the datasets, assisted in the analysis and reviewed the manuscript. NB planned the study, reviewed the analysis, reviewed the manuscript and supervised the project. NB accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Funding This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health (MOH) and the Ministry of Long-Term Care (MLTC). This study also received funding from CIHR. Parts of this material are based on data and information compiled and provided by: MOH, CIHI, CCO, ORG.

Disclaimer The analyses, conclusions, opinions and statements expressed herein are solely those of the authors and do not reflect those of the funding or data sources; no endorsement is intended or should be inferred. Parts of this material are based on data and information provided by Cancer Care Ontario (CCO). The opinions, results, views and conclusions reported in this paper are those of the authors and do not necessarily reflect those of CCO. No endorsement by CCO is intended or should be inferred. Parts of this material are based on data and/or information compiled and provided by CIHI. However, the analyses, conclusions, opinions and statements expressed in the material are those of the author(s), and not necessarily those of CIHI. Parts of this report are based on Ontario Registrar General (ORG) information on deaths, the original source of which is ServiceOntario. The views expressed therein are those of the author and do not necessarily reflect those of ORG or the Ministry of Government Services.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement The dataset from this study is held securely in coded form at ICES. While legal data sharing agreements between ICES and data providers (eg, healthcare organisations and government) prohibit ICES from making the dataset publicly available, access may be granted to those who meet pre-specified criteria for confidential access, available at www.ices.on.ca/DAS (email: das@ices.on.ca). The full dataset creation plan and underlying analytic code are available from the authors upon request, understanding that the computer programs may rely upon coding templates or macros that are unique to ICES and are therefore either inaccessible or may require modification.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC-BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

ORCID iDs
Matthew Castelo http://orcid.org/0000-0001-7477-7400
Nancy N Baxter http://orcid.org/0000-0003-4793-4620

REFERENCES

1. Patel P, De P. Trends in colorectal cancer incidence and related lifestyle risk factors in 15-49-year-olds in Canada, 1969-2010. Cancer Epidemiol 2016;42:90–100.

2. Bailey CE, Hu C-Y, You YN, et al. Increasing disparities in the age-related incidences of colon and rectal cancers in the United States, 1975–2010. JAMA Surg 2015;150:17–22.

3. Crockett E, Buzzoni C, Zappa M. Colorectal cancer incidence rates have decreased in central Italy. Eur J Cancer Prev 2010;19:424–5.

4. Haggar FA, Preen DB, Pereira G, et al. Cancer incidence and mortality trends in Australian adolescents and young adults, 1982–2007. BMC Cancer 2012;12:151.

5. Santucci C, Boffetta P, Levi F, et al. Colorectal cancer mortality in young adults is rising in the United States, Canada, United Kingdom, and Australia but not in Europe and Asia. Gastroenterology 2021;160:1860–2.

6. Cheng E, Blackburn HN, Ng K, et al. Analysis of survival among adults with early-onset colorectal cancer in the National cancer database. JAMA Netw Open 2021;4:e2112539.

7. Gabriell E, Ostapoff K, Atwood K, et al. Disparities in the age-related rates of colorectal cancer in the United States. Am Surg 2017;83:640–7.

8. O’Connell JB, Maggard MA, Livingston EH, et al. Colorectal cancer in the young. Am J Surg 2004;187:343–8.

9. Siegel RL, Jakubowski CD, Fedewa SA. Colorectal cancer in the young: epidemiology, prevention, management. Am Soc Clin Oncol Educ Book Am Soc Clin Oncol Annu Meet 2020;40:1–14.

10. Kim TJ, Kim ER, Hong SN, et al. Long-Term outcome and prognostic factors of sporadic colorectal cancer in young patients: a large Institutional-Based retrospective study. Medicine 2016;95:e3641.

11. Wellar D, Vedsted P, Rubin G, et al. The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. Br J Cancer 2012;106:1262–7.

12. van Erp NF, Heleper CW, Olyhoek SM, et al. Potential for reducing time to referral for colorectal cancer patients in primary care. Ann Fam Med 2019;17:419–27.

13. Delisle M, Helewa RM, Ward MAR, et al. The association between wait times for colorectal cancer treatment and health care costs: a population-based analysis. Dis Colon Rectum 2020;63:165–71.

14. Pearson C, Fraser J, Penna M, et al. Establishing population-based surveillance of diagnostic timeliness using linked cancer registry and administrative data for patients with colorectal and lung cancer. Cancer Epidemiol 2019;61:111–8.

15. Webber C, Flemming JA, Birtwhistle R, et al. Colonoscopy resource availability and its association with the colorectal cancer diagnostic interval: a population-based cross-sectional study. Eur J Cancer Care 2020;29:e13187.

16. ICES. Available: https://www.ices.on.ca/ [Accessed 6 May 2019].

17. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. The Lancet 2007;370:1453–7.
et al. Implementation of ICD-10 in Canada: how has it impacted coded hospital discharge data? *BMC Health Serv Res* 2012;12:1-9.

19 Ontario Cancer Statistics 2020 Data Sources. Cancer care Ontario, 2020. Available: https://www.cancerontario.ca/en/statistical-reports/ontario-cancer-statistics-2020/data-sources [Accessed 21 Mar 2022].

20 Benchimol EI, Guttmann A, Mack DR, et al. Validation of international algorithms to identify adults with inflammatory bowel disease in health administrative data from Ontario, Canada. *J Clin Epidemiol* 2014;67:887–96.

21 Honein-AbouHaidar GN, Baxter NN, Moineddin R, et al. Trends and inequities in colorectal cancer screening participation in Ontario, Canada, 2005-2011. *Cancer Epidemiol* 2013;37:946–56.

22 Walter F, Webster A, Scott S, et al. The Andersen model of total patient delay: a systematic review of its application in cancer diagnosis. *J Health Serv Res Policy* 2012;17:110–8.

23 Groome PA, Webber C, Whitehead M, et al. Determining the cancer diagnostic interval using administrative health care data in a breast cancer cohort. *JCO Clin Cancer Inform* 2019;3:1–10.

24 Flemming JA, Nanji S, Wei X, et al. Association between the time to surgery and survival among patients with colon cancer: a population-based study. *Eur J Surg Oncol* 2017;43:1447–55.

25 Karim S, Paterson B, Kong S. Determining the cancer diagnostic interval using administrative data in a cohort of patients with pancreatic cancer. *JCO* 2021;39:336.

26 Registered Persons Database (RPDB) - Ontario Data Catalogue. Available: https://data.ontario.ca/dataset/registered-persons-database-rpdb [Accessed 31 Dec 2021].

27 Matheson F, van Ingen T. Ontario marginalization index, 2016. Available: https://www.publichealthontario.ca/en/data-and-analysis/health-equity/ontario-marginalization-index.html.

28 Government of Canada SC. Dictionary. Census of Population, 2016 - Economic family after-tax income decile group, 2017. Available: https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/porf/166-eng.cfm [Accessed 31 Dec 2021].

29 FEH J. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer, 2015.

30 Majano SB, Lyratzopoulos G, Rachet B, et al. Do presenting symptoms, use of pre-diagnostic endoscopy and risk of emergency cancer diagnosis vary by comorbidity burden and type in patients with colorectal cancer? *Br J Cancer* 2022;126:652–63.

31 Koo MM, von Wagner C, Abel GA, et al. Typical and atypical presenting symptoms of breast cancer and their associations with diagnostic intervals: evidence from a national audit of cancer diagnosis. *Cancer Epidemiol* 2017;48:140–6.

32 Herbert A, Lyratzopoulos G, Whelan J, et al. Diagnostic timeliness in adolescents and young adults with cancer: a cross-sectional analysis of the BRIGHTLIGHT cohort. *Lancet Child Adolesc Health* 2018;2:188–90.

33 Jensen H, Tørring ML, Olesen F, et al. Diagnostic intervals before and after implementation of cancer patient pathways - a GP survey and registry based comparison of three cohorts of cancer patients. *BMC Cancer* 2015;15:308.

34 Weller D, Monen U, Zaloumina Falborg A, et al. Diagnostic routes and time intervals for patients with colorectal cancer in 10 international jurisdictions; findings from a cross-sectional study from the International cancer benchmarking partnership (ICBP). *BMJ Open* 2018;8:e023870.

35 Bender R, Lange S. Adjusting for multiple testing—when and how? *J Clin Epidemiol* 2001;54:343–9.

36 Di Girolamo C, Walters S, Gildea C, et al. Can we assess cancer waiting time targets with cancer survival? a population-based study of individually linked data from the National cancer waiting times monitoring dataset in England, 2009-2013. *PLoS One* 2018;13:e021288.

37 Wait Times Information System - Introduction. Available: https://ext.cancercare.on.ca/ext/databook/db2021/WTIS/WTIS-Introduction.htm [Accessed 22 Feb 2022].

38 Price S, Spencer A, Zhang X, et al. Trends in time to cancer diagnosis around the period of changing national guidance on referral of symptomatic patients: a serial cross-sectional study using UK electronic healthcare records from 2006-17. *Cancer Epidemiol* 2020;64:101716.

39 Sikkor KC, Dickinson J, Winget M. Factors associated with mode of colorectal cancer detection and time to diagnosis: a population level study. *BMC Health Serv Res* 2017;17:7.

40 Robertson R, Campbell NC, Smith S, et al. Factors influencing time from presentation to treatment of colorectal and breast cancer in urban and rural areas. *Br J Cancer* 2004;90:1479–85.

41 Bergin RJ, Thomas RJF, Whitfield K, et al. Concordance between optimal care pathways and colorectal cancer care: identifying opportunities to improve quality and reduce disparities. *J Eval Clin Pract* 2020;26:918–26.

42 Chen FW, Sundaram V, Chew TA, et al. 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.

43 Neal RD. Do diagnostic delays in cancer matter? *Br J Cancer* 2009;101:Suppl 2:S9–12.

44 Tørring ML, Frydenberg M, Hansen RP, et al. Evidence of increasing mortality with longer diagnostic intervals for five common cancers: a cohort study in primary care. *Eur J Cancer* 2013;49:2187–98.