Predictors of CRC Stage at Diagnosis among Male and Female Adults Participating in a Prospective Cohort Study: Findings from Alberta’s Tomorrow Project

Monica Ghebrial 1, Michelle L. Aktary 2, Qinggang Wang 3, John J. Spinelli 4,5, Lorraine Shack 6, Paula J. Robson 7,8 and Karen A. Kopciuk 3,9,*

Abstract: Colorectal cancer (CRC) is a leading cause of morbidity and mortality in Canada. CRC screening and other factors associated with early-stage disease can improve CRC treatment efficacy and survival. This study examined factors associated with CRC stage at diagnosis among male and female adults using data from a large prospective cohort study in Alberta, Canada. Baseline data were obtained from healthy adults aged 35–69 years participating in Alberta’s Tomorrow Project. Factors associated with CRC stage at diagnosis were evaluated using Partial Proportional Odds models. Analyses were stratified to examine sex-specific associations. A total of 267 participants (128 males and 139 females) developed CRC over the study period. Among participants, 43.0% of males and 43.2% of females were diagnosed with late-stage CRC. Social support, having children, and caffeine intake were predictors of CRC stage at diagnosis among males, while family history of CRC, pregnancy, hysterectomy, menopausal hormone therapy, lifetime number of Pap tests, and household physical activity were predictive of CRC stage at diagnosis among females. These findings highlight the importance of sex differences in susceptibility to advanced CRC diagnosis and can help inform targets for cancer prevention programs to effectively reduce advanced CRC and thus improve survival.

Keywords: colorectal cancer; stage at diagnosis; Alberta’s Tomorrow Project; Canada; social factors; reproductive factors; sex-specific factors

1. Introduction

Colorectal cancer (CRC) is the third most common cancer in Canada and the second leading cause of cancer-attributable deaths [1,2]. Data from Canada [1,3] and other developed nations [4,5] indicate that CRC incidence and mortality are higher among males than females. In Canada, age-standardized incidence rates were estimated at 71.7 per 100,000 for males and 50.9 per 100,000 for females [1]. Similarly, in Alberta, Canada, age-standardized
incidence rates among males and females were estimated at 62.8 per 100,000 and 45.1 per 100,000, respectively [6]. Coinciding with the aging population, it is expected that one in fourteen males and one in sixteen females in Canada will develop CRC in their lifetime [2].

CRC screening programs and early detection reduce CRC incidence and mortality [7–10]. Notably, CRC screening may help diagnose CRC at earlier stages, thereby improving a patient’s chance of survival [10–13]. Cancer staging identifies the severity of cancer at diagnosis according to the size of the tumour, and whether and to what extent it has metastasized to surrounding tissues. Staging ranges from stage I (small and contained) to stage IV (metastasized to other tissues) [10]. In Canada, the net 5-year relative survival of stage IV colon and rectal cancers is estimated at 11% and 12%, respectively, whereas the 5-year survival of stage I colon and rectal cancers is 92% and 87%, respectively [10].

Despite organized CRC screening programs across Canada, in 2015, 49% of CRC diagnoses were stage III and IV [10]. Similarly, in Alberta, 28.6% and 22.0% of all new CRC cases were stage III and IV, respectively [10]. Low population-level participation in screening programs and diagnosis in those who are ineligible for screening may help explain the greater burden of advanced-stage diagnoses.

In addition to low screening participation contributing to later-stage CRC diagnosis, several other important sociodemographic characteristics (e.g., sex, age, race/ethnicity, socioeconomic disadvantage) [10,14–21], and health-related (e.g., family or personal history of polyps or CRC) [14] and psychosocial (e.g., marital status) [15,19,22–25] factors have been shown to influence stage at diagnosis. For instance, while studies have shown higher CRC incidence and mortality among males, some evidence suggests that females are at greater risk of late-stage diagnoses [16,26–28]. The higher risk of advanced CRC among females may be in part due to their greater susceptibility to developing right-sided CRC (i.e., in the proximal colon) [29,30]. Right-sided CRC is less likely to be detected through regular screening methods and typically presents with weaker signs and symptoms (e.g., unrecognized bleeding) than left-sided CRC [29,30]. As a result, right-sided CRC is generally diagnosed at a later stage than left-sided CRC [29,30].

Studies have demonstrated that younger adults are more likely to be diagnosed with late-stage CRC [10,14–16,31]. The higher rates among adults under 50 years may be in part due to CRC screening programs beginning at age 50 years for average-risk individuals [32–34], potentially delaying diagnoses among younger adults [14]. Moreover, younger adults who present with common CRC symptoms such as abdominal pain and blood in the stool may not be first suspected of having CRC, resulting in longer work-up and delayed diagnosis [31]. Individual- and area-level socioeconomic deprivation have also been associated with later-stage CRC [14–16,26]. Advanced CRC rates among low-income populations and ethnic minorities have been observed [16–18,26,35], with these disparities primarily attributed to lower screening participation rates [14,36]. Social support has also been associated with stage of CRC diagnosis. Studies have demonstrated that individuals who were married were more likely to be diagnosed with earlier-stage CRC compared to those who were single, separated, or divorced [22–24]. Individuals who are married are more likely to have higher household income [24] and health insurance [24], both of which are associated with earlier-stage diagnosis [16,18]. Moreover, spouses may encourage and shape health practices [23,37] such as CRC screening participation [24].

Despite the evidence linking several individual- and environmental-level factors to the stage of CRC diagnosis, some studies have failed to find associations [14,15,38,39]. Differences in study designs, health care systems, sample sizes, and model adjustments likely contribute to the mixed findings among studies. In addition, few studies have examined factors associated with CRC stage at diagnosis in Canada [15,35]. Given that stage at diagnosis is a key prognostic factor for CRC survival [14,40], it is crucial to identify factors predisposing males and females to later-stage diagnoses, and how these factors may differ by sex, to enhance treatment effectiveness and survival. Therefore, the objective of this study is to evaluate sociodemographic characteristics, health- and diet-related factors,
and psychosocial factors associated with CRC stage at diagnosis among participants of a large prospective cohort study in Alberta, Canada.

2. Materials and Methods
2.1. Setting and Participants

Alberta’s Tomorrow Project (ATP) is a prospective cohort study examining health, lifestyle, and psychosocial factors associated with the development and prevention of chronic diseases and cancer among healthy adults living in Alberta, Canada [41]. A total of 55,000 adults aged 35–69 years with no reported history of cancer (excluding non-melanoma skin cancer) were enrolled between 2000 and 2015. History of cancer was cross-checked against the Alberta Cancer Registry to confirm self-reports. Participant recruitment was conducted over two phases: first through random digit dialling using Regional Health Authority boundaries as the sampling frame, and then through a combination of recruitment strategies [42]. Further details of participant recruitment and data collection have been described elsewhere [41].

2.2. Data Sources

This study used self-reported data collected from the Health and Lifestyle Questionnaire (HLQ), the Canadian Diet History Questionnaire (CDHQ-I) [43], and the Past-Year Total Physical Activity Questionnaire (PYTPAQ) [44]. Participants completed the surveys at baseline, with a total of 26,890 participants completing all 3 surveys by 2008 [45]. The HLQ assessed sociodemographic characteristics, and psychosocial and health-related variables [46]. The CDHQ-I, a 124-item Food Frequency Questionnaire (FFQ), examined food and nutrient intake over the past year, and the PYTPAQ assessed the type and frequency of physical activity undertaken by participants over the past year [46]. Data from the surveys were subsequently linked to administrative databases using Public Health Numbers for those who consented to data linkage (>99%) [41,42,45].

2.3. Outcome

All CRC cases diagnosed as of 18 January 2018 were included in the analysis. The date and stage of cancer diagnosis were ascertained through data linkage with the Alberta Cancer Registry [41,42,45]. The Alberta Cancer Registry records all individual cancer cases and cancer-related deaths across Alberta [47]. The stage of CRC diagnosis was defined by the Tumour, Node, Metastasis staging system [10].

2.4. Candidate Explanatory Variables

Candidate explanatory variables were selected based on their association with the development of CRC or stage of CRC diagnosis as informed by the literature, including sociodemographic characteristics (e.g., age, sex, total household income), health- (e.g., cancer screening, family history of CRC) and diet-related (e.g., nutrient intake) factors, and psychosocial factors (e.g., stress, social support).

2.5. Statistical Analysis

Descriptive data are presented as median and interquartile range for continuous variables, and as counts and percentages for categorical variables. Differences in variables across CRC stages (i.e., I, II, III/IV) were examined using F-tests for continuous variables and chi-square tests for categorical variables. In a preliminary data assessment, binary variables with low frequencies (i.e., <5) were removed, and categorical variables were collapsed to increase subgroup sizes (i.e., ≥5). Missing responses in continuous variables were imputed using mean value replacement under a missing at random assumption, and missing categorical variables replaced by reference group.

Associations between each candidate explanatory variable and stage at CRC diagnosis were evaluated in bivariate Partial Proportional Odds (PPO) ordinal response models. Variables with a $p$-value < 0.2 were included in the multivariable PPO models. All variables
selected were subsequently evaluated in the multivariable PPO model and were retained if \( p < 0.05 \). Relevant variables associated with the stage at CRC diagnosis were also selected based on the conditional variable importance measure yielded from a random forest analysis using the R package partykit version 1.2-5. Additional variables known to be associated with the stage of CRC diagnosis (i.e., family history of CRC) and those identified from the random forest were forced into the reduced model. The proportional odds assumption was tested, and where the assumption was violated, the model structure was relaxed, allowing the odds ratios (OR) to differ across the outcome cut points (i.e., PPO model).

Functional forms of the continuous variables (e.g., quadratic terms), influence of outliers, and interactions were examined. Model checking was performed using available methods for the binary logistic setting (Stages I and II versus Stages III + IV) in SAS, because many diagnostic tools were not implemented for the PPO model.

Analyses were subsequently stratified to examine whether associations between identified explanatory variables and CRC stage at diagnosis differed by sex. Analyses were conducted in R (version 4.1.0, R Foundation for Statistical Computing, Vienna, Austria) and SAS version 9.4 (SAS Institute, Cary, NC, USA). Statistical significance was set at \( p < 0.05 \).

3. Results

A total of 267 participants (128 males and 139 females) developed CRC over the study period, with a median time of 7 years (quartile 1 to quartile 3: 4 to 10 years) from baseline data collection to receiving a CRC diagnosis. Participant characteristics for males and females by CRC diagnosis stage are presented in Table 1. Among males, 23.4%, 33.6%, and 43.0% were diagnosed with stage I, stage II, and stage III/IV CRC, respectively. The distribution of CRC stage at diagnosis was similar among females, with 22.3%, 34.5%, and 43.2% diagnosed with stage I, stage II, and stage III/IV CRC, respectively. The average age of participants at baseline was 58 years for both males and females, over 96% of participants were Western European ancestry, and a total of 66.2% of females and 64.1% of males had a college or university degree. Just over half (54.7%) of female participants reported a median total annual household income before taxes of <$50,000/year, while a greater proportion of males reported a median total annual household income before taxes of $50,000–$100,000/year. Among males, 76.6% were married and 24.2% were current smokers. Similarly, 71.2% of females were married and 23.7% smoked. Compared to males, females had significantly lower alcohol and red meat consumption, and the higher self-reported CRC screening among females approached significance (\( p = 0.07 \)).

Given the significant differences between males and females in important CRC risk factors, PPO models were stratified by sex. Upon comparing results from the unstratified and stratified analyses, no common factors between sexes were identified. Therefore, only the stratified results are presented. The results from the stratified PPO analysis are summarized in Table 2. Among males, psychosocial factors, including ‘having someone to prepare your meals when unable’ and ‘having someone to hug’ were significantly associated with earlier stage at diagnosis (OR 0.58, 95% CI 0.33–0.99, \( p = 0.05 \) and OR 0.37, 95% CI 0.14–1.01, \( p = 0.05 \), respectively). Likewise, having children was significantly associated with earlier stage at diagnosis when comparing stages III/IV vs. II/I (OR 0.33, 95% CI 0.11–1.00, \( p = 0.05 \)). Total caffeine intake (mg/day) was significantly associated with later stage at diagnosis (OR 1.37, 95% CI 1.06–1.76, \( p = 0.02 \)).
Table 1. Description of the study population by sex and cancer diagnosis stage.

| Characteristic                                      | Male (N = 128) | Female (N = 139) |
|-----------------------------------------------------|----------------|------------------|
| **Age at baseline (years)**                         |                |                  |
| <60                                                 |                |                  |
| 60+                                                 |                |                  |
| **Age at diagnosis (years)**                        |                |                  |
| <60                                                 |                |                  |
| 60+                                                 |                |                  |
| **Education**                                       |                |                  |
| High school diploma                                |                |                  |
| Post-secondary                                     |                |                  |
| **Married or common in-law**                        |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Employment**                                      |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Household Income ($)**                            |                |                  |
| <50 K                                               |                |                  |
| 50–100 K                                           |                |                  |
| >100 K                                              |                |                  |
| **Geography residence**                             |                |                  |
| Rural                                               |                |                  |
| Urban                                               |                |                  |
| **First degree colorectal cancer family history**    |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Ever had a digital rectal exam**                  |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Ever had a blood stool test**                     |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Have ever had a sigmoidoscopy or colonoscopy done**|                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Have ever had a PSA blood test**                  |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Have ever been pregnant**                         |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Have ever had a hysterectomy**                    |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Female Hormones used for menopause**              |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Have ever had Pap smear test**                    |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
| **Have ever had mammogram test**                    |                |                  |
| Yes                                                 |                |                  |
| No                                                  |                |                  |
### Table 1. Cont.

| Characteristic | Male (N = 128) | Female (N = 139) |
|----------------|---------------|------------------|
| **Type of smoker** |               |                  |
| Non smoker | 10 (7.8) | 14 (10.9) |
| past smoker | 16 (12.5) | 18 (14.1) | 28 (21.9) | 62 (46.4) | 13 (9.4) | 21 (15.1) |
| current smoker | 4 (3.1) | 11 (8.6) | 16 (12.5) | 31 (24.2) | 6 (4.3) | 15 (10.8) | 12 (8.6) | 33 (23.7) |

| Spend in the sun 11 a.m.–4 p.m. in June–August (hours) | Male (N = 128) | Female (N = 139) |
|--------------------------------------------------------|---------------|------------------|
| <1 | 23 (18.0) | 33 (25.8) | 37 (28.9) | 93 (72.7) | 15 (10.8) | 27 (19.4) | 31 (22.3) | 73 (52.5) |
| ≥1 | 7 (5.5) | 10 (7.8) | 18 (14.1) | 35 (27.3) | 6 (4.3) | 15 (10.8) | 12 (8.6) | 33 (23.7) |

| Any current stressful situations | Male (N = 128) | Female (N = 139) |
|---------------------------------|---------------|------------------|
| None | 19 (14.8) | 27 (21.1) | 29 (22.7) | 75 (58.6) | 19 (13.7) | 23 (16.6) | 40 (28.8) | 82 (59.0) |
| ≥1 | 11 (8.6) | 16 (12.5) | 26 (20.3) | 53 (41.4) | 12 (8.6) | 25 (18.0) | 20 (14.4) | 57 (41.0) |

| Any children | Male (N = 128) | Female (N = 139) |
|---------------|---------------|------------------|
| Yes | 26 (20.3) | 40 (31.3) | 44 (34.4) | 110 (85.9) | 29 (20.9) | 46 (33.1) | 53 (38.1) | 128 (92.1) |
| No | 4 (3.1) | 3 (2.3) | 11 (8.6) | 18 (14.1) | 2 (1.4) | 2 (1.4) | 7 (5.0) | 11 (7.9) |

| Too much is expected of you by others | Male (N = 128) | Female (N = 139) |
|--------------------------------------|---------------|------------------|
| Yes | 9 (7.0) | 12 (9.4) | 17 (13.3) | 38 (29.7) | 12 (8.6) | 18 (13.0) | 15 (10.8) | 45 (32.4) |
| No | 21 (16.4) | 31 (24.2) | 38 (29.7) | 90 (70.3) | 19 (13.7) | 30 (21.6) | 45 (32.4) | 94 (67.6) |

### Mean (SD)

| Characteristic | Male | Female |
|----------------|------|--------|
| Age at baseline (years) | 56.3 (7.8) | 58.3 (7.1) |
| Distance to accessible health centre by a vehicle (minutes) | 11.6 (11.8) | 13.1 (12.3) |
| Total body mass index (BMI) | 28.9 (5.6) | 30.0 (5.1) |
| Total physical activity (Met-hours/week) | 164.5 (81.5) | 136.3 (74.1) |
| Total recreational activity (Met-hours/week) | 34.8 (24.8) | 26.9 (20.7) |
| Total household activity (Met-hours/week) | 35.1 (29.1) | 38.3 (33.7) |
| Total dietary caloric intake (kcal/day) | 2239 (1035) | 2235 (980) |
| Total protein intake (g/day) | 87 (39) | 79 (31) |
| Total caffeine intake (mg/day) | 89 (21) | 83 (21) |
| Social support: someone to hug: scale 1–5 | 5 (5–5) | 4 (3–5) |
| Social support: someone to prepare your meals when you were unable to do it yourself: scale 1–5 | 5 (5–5) | 4 (3–5) |
| Total alcohol intake (g/day) | 13 (2–23) | 5 (1–15) |
| Daily fruit and vegetable intake (servings/day) | 4 (3–6) | 3 (2–6) |
| Daily red meat intake (servings/day) | 3 (2–4) | 2 (1–3) |

| Statistical significance within each gender at *p < 0.05, **p < 0.01. £ These variables are derived from the Alberta Cancer Registry, Administration data. Elixhauser Comorbidity Index: range (0–7). $ indicates number for age was <10 so cell entry suppressed. |
Table 2. Predictors of colorectal cancer diagnosis at different stages from the Partial Proportional Odds Regression Model.

| Predictor Variable                          | Males (N = 128) | Females (N = 139) |
|--------------------------------------------|-----------------|-------------------|
|                                            | Cancer Stage (PPO) ORs (95% CI) | p Value | Cancer Stage (PPO) ORs (95% CI) | p Value |
| Colorectal cancer family history           |                 |                   |                                |         |
| Yes vs. No                                 | 1.24 (0.48–3.24) | 0.66              | III/IV vs. II & I              | 1.65 (0.58–4.74) | 0.35 |
|                                            |                 |                   | III/IV & II vs. I              | 0.21 (0.07–0.66) | 0.01 |
| Have ever been pregnant                    |                 |                   |                                |         |
| Yes vs. No                                 | 0.21 (0.05–0.87) | 0.03              |                                |         |
| Have ever had a hysterectomy               |                 |                   |                                |         |
| Yes vs. No                                 | III/IV vs. II & I | 0.58 (0.26–1.33) | 0.20                           |         |
|                                            | III/IV & II vs. I | 0.28 (0.11–0.74) | 0.01                           |         |
| Female Hormones used for menopause         |                 |                   |                                |         |
| Yes vs. No                                 | III/IV vs. II & I | 3.04 (1.41–6.58) | <0.01                         |         |
|                                            | III/IV & II vs. I | 0.69 (0.28–1.73) | 0.43                           |         |
| Number of Pap tests in lifetime            |                 |                   |                                |         |
|                                            | 0.96 (0.93–1.00) | 0.04              |                                |         |
| Total household activity (Met-hours/week)  |                 |                   |                                |         |
| <40                                        | 1.01 (1.00–1.01) | 0.05              |                                |         |
| 40–60                                      | 4.08 (1.54–10.82)| <0.01             | 3.53 (1.49–8.37) | <0.01 |
| >60                                        | 5.33 (2.58–11.01)| <0.01             |                                |         |
| Any children                               |                 |                   |                                |         |
| Yes vs. No                                 | III/IV vs. II & I | 0.33 (0.11–1.00) | 0.05                           |         |
|                                            | III/IV & II vs. I | 1.57 (0.39–6.33) | 0.53                           |         |
| Social support: someone to hug: scale 1–5  |                 |                   |                                |         |
|                                            | III/IV vs. II & I | 1.56 (0.90–2.71) | 0.12                           |         |
|                                            | III/IV & II vs. I | 0.37 (0.14–1.01) | 0.05                           |         |
| Social support: someone to prepare your meals when you were unable to do it yourself: scale 1–5 | | | | |
|                                            | III/IV vs. II & I | 0.58 (0.33–0.99) | 0.05                           |         |
|                                            | III/IV & II vs. I | 0.90 (0.40–2.01) | 0.80                           |         |
| Total caffeine intake (mg/day) at log scale|                 |                   |                                |         |
|                                            | III/IV vs. II & I | 1.22 (0.92–1.63) | 0.18                           |         |
|                                            | III/IV & II vs. I | 1.37 (1.06–1.76) | 0.02                           |         |

Note that: Cancer stage III/IV means stage III & IV. Whenever there is entry for the cancer stage column, it means the results are from a partial proportional odds model. Otherwise, it is the proportional odds model where coefficients are same for stages III/IV vs. II & I and III/IV & II vs. I.

Among females, health-related factors that were associated with earlier stage at diagnosis when comparing stages III/IV & II vs. I included family history of CRC (OR 0.21, 95% CI 0.11–0.74, p = 0.01) and previously having a hysterectomy (OR 0.28, 95% CI 0.11–0.74, p = 0.01). Number of Pap tests in a lifetime (OR 0.96, 95% CI 0.60–1.51, p = 0.05) and history of pregnancy (OR 0.21, 95% CI 0.05–0.87, p = 0.03) were also protective against late-stage diagnosis. However, history of menopausal hormone therapy (MHT) was significantly associated with later-stage diagnosis (stages III/IV vs. II/I) (OR 3.04, 95% CI 1.41–6.58, p < 0.01). In addition, total household physical activity was significantly associated with later-stage CRC diagnosis (OR 1.01, 95% CI 1.00–1.01, p = 0.05).

4. Discussion
Cancer stage at diagnosis is an important indicator of treatment effectiveness and survival [10]. Despite organized screening programs across Canada and universal health care coverage, nearly 50% of CRC cases are diagnosed at a late stage (III or IV) [10].
This study examined associations between CRC stage at diagnosis and sociodemographic characteristics, health- and diet-related factors, and psychosocial factors among male and female participants in a prospective cohort study in Alberta, Canada. From 2000 to 2018, 267 participants in ATP were diagnosed with CRC, 43.1% of which were late-stage diagnoses (stage III/IV). The overall distribution of CRC by stage was similar among male and female participants. Whereas some studies have demonstrated that females are more frequently diagnosed at a later stage [16,26–28], our findings are comparable to Canadian statistics [10].

The current study found that psychosocial and diet-related factors were predictors of CRC stage at diagnosis among males, while health-related factors were more predictive of CRC stage at diagnosis among females. Among males, social support was inversely associated with CRC stage at diagnosis. Evidence suggests that social support is an important factor in CRC outcomes, including the stage at diagnosis [24,48]. A recent systematic review of social factors associated with CRC outcomes identified a number of studies that showed earlier-stage CRC diagnosis and improved CRC survival among those who were married or living with a partner compared to their single counterparts [24]. The inverse association between social support and CRC stage observed only in males may be partly explained by marital status, as studies have shown that males report more social support from their spouse, while females report drawing more social support from friends and relatives [49]. Furthermore, females are more likely to monitor their spouse’s health practices (e.g., cancer screening, attending physician appointments, and dietary intake) and overall health [50,51].

Having children was also inversely associated with CRC stage at diagnosis among males. The protective effects of parenthood against later-stage CRC may be attributed to higher levels of social support, as having children is closely tied to marital status [52,53]. For instance, Kravdal [52] found that the mortality rate of married males with children was 30% lower than that of males without children and those who were never married. Parenthood may also improve CRC outcomes, as having young children is associated with having a stay-at-home partner (generally the mother [53]) who can aid in scheduling healthcare appointments [23]. However, studies have also demonstrated that having children is protective against cancer mortality, even after controlling for marital status [52]. Thus, the link between having children and reduced cancer risk and mortality may be explained by other protective effects conferred by parenthood. For instance, having children may promote feelings of being valued and connected to others and, thus, encourage more self-care behaviours, including participating in cancer screening [48].

This study found diet-related factors associated with the stage of CRC diagnosis among males. Specifically, caffeine intake was associated with increased risk of developing later-stage CRC. Whereas some studies have demonstrated protective effects of coffee [54–57], tea [58,59], and caffeine [57] consumption against CRC risk and mortality, others have reported null [56,60,61] or harmful effects on CRC outcomes [62,63]. The role of coffee and caffeine in reducing CRC risk may be through various mechanisms, such as anti-proliferative, antioxidant, and anti-inflammatory effects [55,56,62]. However, coffee also contains compounds with mutagenic effects (e.g., glyoxal [64]) that may induce carcinogenesis [56]. Moreover, some studies have identified significant inverse associations between decaffeinated coffee and CRC risk [54,62], suggesting that the observed benefits of coffee consumption may be attributable to other compounds found in coffee rather than caffeine alone. Some studies have suggested that associations between coffee and caffeine consumption and CRC risk may be sex-specific [65]. Higher coffee consumption is associated with the male sex, along with alcohol consumption [60], smoking, and poorer dietary patterns [62,66]. Therefore, the association between caffeine intake and advanced CRC stage at diagnosis in males observed in this study may be due to residual confounding [67,68].

This study identified important health-related factors associated with stage of CRC diagnosis in females. MHT was associated with increased risk of later stage CRC, while history of pregnancy or hysterectomy, lifetime number of Pap tests (cervical cancer screen-
ing), and family history of CRC were associated with decreased risk of late-stage CRC. The significant association between MHT and the increased risk of later-stage CRC identified in this study aligns with findings from a randomized controlled trial demonstrating that women on MHT who developed CRC were diagnosed at a later stage than those on the placebo [69]. Whereas a small number of studies have demonstrated associations between MHT usage and later-stage CRC [69,70] and CRC risk and mortality [71,72], study findings have been inconsistent [70,73–82].

The protective effects of MHT against CRC among females can be partly attributed to the role of oestrogen and the expression of intracellular oestrogen receptors (ER) on inhibiting CRC cell growth [83,84]. Two predominant ER involved in colorectal tumorigenesis include ER-beta and ER-alpha, and their expression may be influenced by timing of MHT initiation. In the pre-cancerous stages of CRC development, exogenous oestrogen activates ER-beta, inhibiting the growth of cancer cells and stimulating apoptosis [85]. However, in the later stages of CRC, the expression of ER-beta is reduced and ER-alpha increases [85]. Exogenous oestrogen then activates ER-alpha, promoting the metastasis of CRC cells [85]. Therefore, it is possible that females in this study began MHT in the later stages of CRC initiation and progression, increasing the risk of advanced CRC.

History of pregnancy and hysterectomy were associated with reduced risk of late-stage CRC diagnosis. Several studies [52,76,86], but not all [81,87–90], have shown reductions in CRC risk and mortality in females with previous pregnancies compared to their nulliparous counterparts. The protective effects of pregnancy may be due to the higher levels of oestradiol over the course of the pregnancy, and the subsequent effect of oestrogen on reducing bile acid synthesis and insulin-like growth factors [91,92], resulting in reduced risk of carcinogenesis [72,77]. Moreover, investigations of the associations between hysterectomy and CRC risk have either found no associations [78,86] or an increase in CRC risk [92]. A meta-analysis showed a 24% increased risk of CRC for females who had undergone a hysterectomy compared to those who had not undergone surgery [92]. The reduced oestrogen and progesterone levels following a hysterectomy are thought to be involved in increasing CRC risk [92,93].

Associations between reproductive factors and CRC stage at diagnosis may also be related to sociocultural factors. Females are more likely than males to regularly visit their physician, even after adjusting for reproductive health care visits [94,95]. Indeed, Canadian data suggest that 85.5% of Canadians (85.1% of Albertans) have a regular family physician, with physician attachment higher among females than males [96]. Regular physician visits, increased exposure to cancer screening, and physician-made CRC screening recommendations have shown to improve CRC screening uptake and participation [94,97–101]. Moreover, family history of CRC is an important risk factor for CRC [14]. As a result, females with a family history of CRC may be more likely to undergo CRC screening [102].

Higher levels of household physical activity (e.g., housework) were associated with later stage CRC diagnosis among females. Contrary to our findings, evidence has demonstrated that higher levels of physical activity are associated with reduced CRC risk [61,103–106]. Moreover, a recent meta-analysis found a significant inverse association between household physical activity and CRC risk among females (relative risk (RR) 0.78, 95% CI 0.69–0.88) but not males (RR 1.04, 95% CI 0.84–1.30) [105]. However, studies have demonstrated that females tend to spend more time in occupational and household activities than recreational activities [106–110], which may create time barriers and thus potentially delay physician visits or cancer screening.

Our study has several strengths. First, this study examined associations between several important sociodemographic characteristics, health- and diet-related factors, and psychosocial factors and CRC stage at diagnosis using data from a large population-level cohort. Second, the prospective cohort study design allowed the assessment of CRC risk factors prior to cancer diagnosis, thus minimizing recall bias [111]. Third, we collected data using valid and reliable surveys and assessment tools [44,112]. However, the study results should be interpreted considering some limitations. Data for this study were
collected using self-reported measures, and thus are subject to social desirability bias [113]. Social desirability bias may be of particular concern with measures such as self-reported physical activity [113,114], dietary intake [115], and less socially acceptable practices, such as smoking and excess alcohol consumption [116]. In addition, dietary intake was examined using an FFQ. Whereas FFQs are recommended dietary assessment tools, they are subject to recall bias and measurement error [117]. Our study had limited ethnic/racial diversity, with over 90% of participants of European descent, thus limiting generalizability of these results. However, according to statistics from 2001 (when these ATP data were being collected), close to 90% of individuals living in Alberta were White [118]. Therefore, the study sample is representative of the Alberta population in the early 2000s. Moreover, the surveys did not allow for detailed assessment of MHT. Given the complexity of hormonal and physiological processes, the ability to stratify our analysis by MHT type, timing, or duration may have provided further insight into the observed association. However, with only 139 cases of CRC detected in females over the study period, our study was not powered to further stratify our analyses. Finally, while our analyses controlled for important measured confounders, we cannot rule out the possibility of residual confounding.

5. Conclusions

This study found several novel factors associated with CRC stage at diagnosis among males and females. Social support and caffeine intake were important predictors of CRC stage among males, while reproductive and hormonal factors and household physical activity predicted CRC stage at diagnosis among females. These findings highlight important sex-specific differences in susceptibility to advanced CRC. Whereas regular CRC screening remains fundamental in the detection of earlier-stage CRC, identifying other important factors associated with CRC stage at diagnosis, particularly those specific to males and females, can help inform targets for cancer prevention programs to effectively reduce advanced CRC and thus greatly improve CRC survival.

Author Contributions: M.L.A., M.G. and K.A.K. wrote the manuscript; Q.W. conducted the analysis; K.A.K., P.J.R. and L.S. obtained funding; and all authors (M.G., M.L.A., Q.W., J.J.S., L.S., P.J.R. and K.A.K.) contributed to study design and read and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Canadian Institutes for Health Research (Primary Investigators Kopciuk, Robson, and Shack), grant number PJT-148774.

Institutional Review Board Statement: Ethics approval for this study (HREBA.CC-16-0814) and ATP (HREBA.CC-17-0461) was granted from the Health Research Ethics Board of Alberta-Cancer Committee.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Access to individual-level data is available in accordance with the Health Information Act of Alberta and Alberta’s Tomorrow Project (ATP). Access Guidelines at https://myatpresearch.ca (accessed on 18 January 2018).

Acknowledgments: Alberta’s Tomorrow Project is only possible because of the commitment of its research participants, its staff, and its funders: Alberta Health, Alberta Cancer Foundation, Canadian Partnership Against Cancer and Health Canada, and substantial in-kind funding from Alberta Health Services. Cancer registry data were obtained through linkage with Surveillance & Reporting, Cancer Research & Analytics, Cancer Care Alberta. The views expressed herein represent the views of the author(s) and not of Alberta’s Tomorrow Project or any of its funders. This study is based in part on data provided by Alberta Health. The interpretation and conclusions contained herein are those of the researchers and do not necessarily represent the views of the Government of Alberta. Neither the Government of Alberta nor Alberta Health express any opinion in relation to this study.

Conflicts of Interest: The authors declare no conflict of interest or competing interest.
77. McMichael, A.J.; Potter, J.D. Reproduction, Endogenous and Exogenous Sex Hormones, and Colon Cancer—A Review and Hypothesis. *J. Natl. Cancer Inst.* 1980, 65, 1201–1207. [PubMed]

78. Kampman, E.; Potter, J.D.; Slattery, M.L.; Caan, B.J.; Edwards, S. Hormone Replacement Therapy, Reproductive History, and Colon Cancer: A multicenter, Case-Control Study in the United States. *Cancer Causes Control* 1997, 8, 146–158. [CrossRef]

79. Terry, M.B.; Neugut, A.I.; Bostick, R.M.; Sandler, R.S.; Haile, R.W.; Jacobson, J.S.; Fenoglio-Preiser, C.M.; Potter, J.D. Risk Factors for Advanced Colorectal Adenomas: A Pooled Analysis. *Cancer Epidemiol. Biomark. Prev.* 2002, 11, 622–629.

80. Nakhoostin, L.; Ladstler, A.; Stute, P. Impact of Menopausal Hormone Therapy on Colorectal Cancer Risk—A Systematic Review. *Clin. Endocrinol.* (Oxf) 2021, 95, 390–397. [CrossRef]

81. Kabat, G.C.; Miller, A.B.; Rohan, T.E. Oral Contraceptive Use, Hormone Replacement Therapy, Reproductive History and Risk of Colorectal Cancer in Women. *Int. J. Cancer* 2008, 122, 643–646. [CrossRef] [PubMed]

82. Ritenbaugh, C.; Stanford, J.L.; Wu, L.; Shikany, J.M.; Taylor, V.; Garland, C.; Frank, G.; Lane, D.; et al. Conjugated Equine Estrogens and Colorectal Cancer Incidence and Survival: The Women’s Health Initiative Randomized Clinical Trial. *Cancer Epidemiol. Biomark. Prev.* 2008, 17, 2609–2618. [CrossRef] [PubMed]

83. Papaxoinis, K.; Triantafyllou, K.; Sasco, A.J.; Nicolopoulos-Stamati, P.; Ladas, S.D. Subsite-Specific Differences of Estrogen Receptor Beta Expression in the Normal Colon Epithelium: Implications for Carcinogenesis and Colorectal Cancer Epidemiology. *Eur. J. Gastroenterol. Hepatol.* 2010, 22, 614–619. [CrossRef] [PubMed]

84. Barzi, A.; Lenz, A.M.; Labonte, M.J.; Lenz, H.J. Molecular Pathways: Estrogen Pathway in Colorectal Cancer. *Clin. Cancer Res.* 2013, 19, 5842–5848. [CrossRef] [PubMed]

85. Chen, J.; Iverson, D. Estrogen in Obesity-Associated Colon Cancer: Friend or Foe? Protecting Postmenopausal Women but Promoting Late-Stage Colon Cancer. *Cancer Control* 2012, 19, 1767–1773. [CrossRef] [PubMed]

86. Murphy, N.; Xu, L.; Zervoudakis, A.; Xue, X.; Kabat, G.; Rohan, T.E.; Wassertheil-Smoller, S.; O’Sullivan, M.J.; Thomson, C.; Messina, C.; et al. Reproductive and Menstrual Factors and Colorectal Cancer Incidence in the NIH-AARP Diet and Health Study. *Br. J. Cancer* 2017, 116, 117–125. [CrossRef]

87. Arem, H.; Park, Y.; Felix, A.S.; Zervoudakis, A.; Brinton, L.A.; Matthews, C.E.; Gunter, M.J. Reproductive and Hormonal Factors and Mortality among Women with Colorectal Cancer in the NIH-AARP Diet and Health Study. *Br. J. Cancer* 2015, 113, 562–568. [CrossRef] [PubMed]

88. Martinez, M.E.; Grodstein, F.; Giovannucci, E.; Colditz, G.A.; Speizer, F.E.; Hennekens, C.; Rosner, B.; Willett, W.C.; Stampfer, M.J. A Prospective Study of Reproductive Factors, Oral Contraceptive Use, and Risk of Colorectal Cancer. *Cancer Epidemiol. Biomark.* Prev. 1997, 6, 1–5.

89. Jacobsen, B.K.; Vollset, S.E.; Kvale, G. Do Reproductive Factors Influence Colorectal Cancer Survival? *J. Clin. Epidemiol.* 1995, 48, 1119–1122. [CrossRef]

90. Zervoudakis, A.; Strickler, H.D.; Park, Y.; Xue, X.; Hollenbeck, A.; Schatzkin, A.; Gunter, M.J. Reproductive History and Risk of Colorectal Cancer in Postmenopausal Women. *J. Natl. Cancer Inst.* 2011, 103, 826–834. [CrossRef] [PubMed]

91. Wernli, K.J.; Wang, Y.; Zheng, Y.; Potter, J.D.; Newcomb, P.A. The Relationship between Gravidity and Parity and Colorectal Cancer Risk. *J. Women’s Health* 2009, 18, 995–1001. [CrossRef] [PubMed]

92. Luo, G.; Zhang, Y.; Wang, L.; Huang, Y.; Yu, Q.; Guo, P.; Li, K. Risk of Colorectal Cancer with Hysterectomy and Oophorectomy: A Systematic Review and Meta-Analysis. *Int. J. Surg.* 2016, 34, 88–95. [CrossRef]

93. Segelman, J.; Lindstrom, L.; Frisell, J.; Lu, Y. Population-Based Analysis of Colorectal Cancer Risk after Oophorectomy. *Br. J. Surg.* 2016, 103, 908–915. [CrossRef] [PubMed]

94. Ritvo, P.; Myers, R.E.; Paszat, L.; Serenity, M.; Perez, D.F.; Rabeneck, L. Gender Differences in Attitudes Impeding Colorectal Cancer Screening. *BMC Public Health* 2013, 13, 500. [CrossRef] [PubMed]

95. Courtney, W.H. Constructions of Masculinity and Their Influence on Men’s Well-Being—a Theory of Gender and Health. *Soc. Sci. Med.* 2000, 50, 1385–1401. [CrossRef]

96. Statistics Canada. Primary Health Care Providers. 2019. Available online: https://www150.statcan.gc.ca/n1/pub/82-625-x/20001/article/0004-eng.htm (accessed on 14 May 2021).

97. Brenes, G.A.; Paskett, E.D. Predictors of Stage of Adoption for Colorectal Cancer Screening. *Prev. Med.* 2000, 31, 410–416. [CrossRef]

98. Lipkus, I.M.; Rimer, B.K.; Lyons, P.R.; Pradhan, A.A.; Conaway, M.; Woods-Powell, C.T. Colorectal Screening Patterns and Perceptions of Risk among African-American Users of a Community Health Center. *J. Community Health* 2009, 34, 995–1001. [CrossRef]

99. Menees, S.B.; Inadomi, J.; Elta, G.; Korsnes, S.; Punch, M.; Aldrich, L. Colorectal Cancer Screening Compliance and Contemplation in Gynecology Patients. *J. Women’s Health* 2010, 19, 911–917. [CrossRef]

100. Carlos, R.C.; Fendrick, A.M.; Ellis, J.; Bernstein, S.J. Can Breast and Cervical Cancer Screening Visits Be Used to Enhance Colorectal Cancer Screening? *J. Am. Coll. Radiol.* 2004, 1, 769–776. [CrossRef]

101. McGregor, S.E.; Bryant, H.E. Predictors of Colorectal Cancer Screening: A Comparison of Men and Women. *Can. J. Gastroenterol.* 2005, 19, 334–349. [CrossRef]

102. Beydoun, H.A.; Beydoun, M.A. Predictors of Colorectal Cancer Screening Behaviors Among Average-Risk Older Adults in the United States. *Cancer Causes Control* 2008, 19, 339–359. [CrossRef]
103. Rezende, L.F.M.; Sa, T.H.; Markozannes, G.; Rey-Lopez, J.P.; Lee, I.M.; Tsilidis, K.K.; Ioannidis, J.P.A.; Eluf-Neto, J. Physical Activity and Cancer: An Umbrella Review of the Literature Including 22 Major Anatomical Sites and 770,000 Cancer Cases. *Br. J. Sports Med.* 2018, 52, 826–833. [CrossRef] [PubMed]

104. Boyle, T.; Keegel, T.; Bull, F.; Heyworth, J.; Fritschi, L. Physical Activity and Risks of Proximal and Distal Colon Cancers: A Systematic Review and Meta-Analysis. *J. Natl. Cancer Inst.* 2012, 104, 1548–1561. [CrossRef] [PubMed]

105. Mahmood, S.; MacInnis, R.J.; English, D.R.; Karahalios, A.; Lynch, B.M. Domain-Specific Physical Activity and Sedentary Behaviour in Relation to Colon and Rectal Cancer Risk: A Systematic Review and Meta-Analysis. *Int. J. Epidemiol.* 2017, 46, 1797–1813. [CrossRef] [PubMed]

106. Shi, Y.; Li, T.; Wang, Y.; Zhou, L.; Qin, Q.; Yin, J.; Wei, S.; Liu, L.; Nie, S. Household Physical Activity and Cancer Risk: A Systematic Review and Dose-Response Meta-Analysis of Epidemiological Studies. *Sci. Rep.* 2015, 5, 14901. [CrossRef] [PubMed]

107. Ainsworth, B.E. Issues in the Assessment of Physical Activity in Women. *Res. Q. Exerc. Sport* 2000, 71 (Suppl. 2), 37–42. [CrossRef] [PubMed]

108. Phongsavan, P.; Merom, D.; Marshall, A.; Bauman, A. Estimating Physical Activity Level: The Role of Domestic Activities. *J. Epidemiol. Community Health* 2004, 58, 466–467. [CrossRef]

109. Murphy, M.H.; Donnelly, P.; Breslin, G.; Shi, B.; Nevill, A.M. Does Doing Housework Keep You Healthy? The Contribution of Domestic Physical Activity to Meeting Current Recommendations for Health. *BMC Public Health* 2013, 13, 966. [CrossRef]

110. Schor, J.B. *The Overworked American*; Basic Books; Reprint edition; Kindle Edition: New York, NY, USA, 1992.

111. Lee, W.; Hotopf, M. Chapter 10—Critical Appraisal: Reviewing Scientific Evidence and Reading Academic Papers. In *Core Psychiatry*; WB Saunders: Philadelphia, PA, USA, 2012; pp. 131–142. [CrossRef]

112. Subar, A.F.; Thompson, F.E.; Kipnis, V.; Midhune, D.; Hurwitz, P.; McNutt, S.; Mcintosh, A.; Rosenfeld, S. Comparative Validation of the Block, Willett, and National Cancer Institute Food Frequency Questionnaires-The Eating at America’s Table Study. *Am. J. Epidemiol.* 2001, 154, 1089–1099. [CrossRef] [PubMed]

113. Adams, S.A.; Matthews, C.E.; Ebbeling, C.B.; Moore, C.G.; Cunningham, J.E.; Fulton, J.; Hebert, J.R. The Effect of Social Desirability and Social Approval on Self-Reports of Physical Activity. *Am. J. Epidemiol.* 2005, 161, 389–398. [CrossRef]

114. Du, H.; Li, L.; Whitlock, G.; Bennett, D.; Guo, Y.; Bian, Z.; Chen, J.; Sherliker, P.; Huang, Y.; Zhang, N.; et al. Patterns and Socio-demographic Correlates of Domain-Specific Physical Activities and Their Associations with Adiposity in the China Kadoorie Biobank Study. *BMC Public Health* 2014, 14, 826. [CrossRef]

115. Hebert, J.R.; Clemow, L.; Pbert, L.; Ocken, I.S.; Ockene, J.K. Social Desirability Bias in Dietary Self-Report May Compromise the Validity of Dietary Intake Measures. *Int. J. Epidemiol.* 1995, 24, 389–398. [CrossRef] [PubMed]

116. Viner, B.; Barberio, A.M.; Haig, T.R.; Friedenreich, C.M.; Brenner, D.R. The Individual and Combined Effects of Alcohol Consumption and Cigarette Smoking on Site-Specific Cancer Risk in a Prospective Cohort of 26,607 adults: Results from Alberta’s Tomorrow Project. *Cancer Causes Control* 2019, 30, 1313–1326. [CrossRef] [PubMed]

117. National Cancer Institute. Dietary Assessment Primer. Available online: https://dietassessmentprimer.cancer.gov/profiles/record/ (accessed on 21 April 2021).

118. Government of Alberta. *Demographic Spotlight-The Visible Minority Population: Recent Trends in Alberta and Canada*; Finance and Enterprise: Edmonton, AB, Canada, 2011.