Association Between Prediagnostic Cigarette Smoking and Colorectal Cancer Survival by Molecular Subtypes and Age-onset Status

Zacharie Eric Lolema II Bakwa

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2023

Committee:

Amanda I. Phipps

Margaret Madeleine

William M. Grady

Program Authorized to Offer Degree:

Epidemiology

University of Washington

Abstract

Association Between Prediagnostic Cigarette Smoking and Colorectal Cancer Survival by Molecular Subtypes and Age-onset Status

Zacharie Eric Lolema II Bakwa

Chair of the Supervisory Committee:
Amanda I. Phipps
Department of Epidemiology

Background: In the United States, colorectal cancer (CRC) remains the third most commonly diagnosed malignancy in both men and women, and approximately 36% of CRC patients die within the first five years after diagnosis. Several factors are known to affect the survival of CRC patients, including smoking history; however, it is unclear whether previously-observed associations between cigarette smoking history and CRC survival vary across molecular subtypes of CRC or by age at CRC diagnosis**.**

Methods: We conducted a retrospective analysis of survival using data from two complementary study populations: the Colon Cancer Family Registry (CCFR) and the Advanced Colorectal Cancer of Serrated Subtype (ACCESS). This analysis included 4,901 participants from the CCFR and ACCESS studies who completed a baseline questionnaire at the time of enrollment. We used Pearson Chi-square tests of independence to determine if there was a statistically significant difference in the distribution of cigarette smoking status by age at diagnosis as a categorical variable (<50 years and ≥ 50 years) and by different molecular

subtypes of CRC. We used univariate and multivariate Cox Proportional Hazards regression analyses to describe associations of smoking with CRC survival overall and within groups defined by age at diagnosis and molecular subtypes.

Results: Of the 4,901 participants included in this study, 1,378 participants died due to CRC during study follow-up (average follow-up period = 13.6 years). Ever smoker participants had 27% higher instantaneous risk of dying from CRC than never-smokers (HR=1.27; 95% CI:1.13-1.42). Even after adjusting for different covariates, ever smoker participants continued to have 21% higher risk of dying from CRC compared to never-smoker patients (HR=1.21; 95% CI:1.06-1.38). In analyses stratified by age at diagnosis, the observed association with smoking was limited to individuals with later-onset CRC. In analyses stratified by tumor molecular subtypes, a statistically significant association of smoking with CRC survival was observed only among participants with CRC type 4 (CIMP-low or negative, MSI-low or MSS, *BRAF* and *KRAS* wildtype), although similar patterns of association were noted for most other subtypes.

Conclusion: Cigarette smoking is associated with poorer survival after CRC diagnosis, particularly for individuals diagnosed with CRC at later ages.

Introduction

In the United States, colorectal cancer (CRC) remains the third most frequently diagnosed malignancy in both men and women, and approximately 36% of CRC patients die within the first five years after diagnosis.[1,2] Several factors are known to influence survival in patients with CRC, including the stage of the tumor at the time of diagnosis, tumor location, age at diagnosis, access to treatment, and low socioeconomic status, to name a few.[2-5] Molecular characteristics of CRC are also increasingly recognized as being associated with survival differences among CRC patients.[6-9] In particular, the presence (or absence) of microsatellite instability (MSI) status, CpG island methylator phenotype (CIMP) status, and somatic hotspot mutations in *KRAS* and *BRAF* have been found to be associated with differences in patient survival.[6-9] Authors have proposed classification of CRC by combining these common biomarkers to define molecular subtypes of CRC, and these molecular subtypes have also been associated with differences in survival among patients with CRC.[10,11]

Cigarette smoking is an established risk factor for the development of CRC and has also been associated with CRC survival[7-9,11] and with certain individual tumor characteristics which are themselves associated with survival (e.g., MSI status, CIMP status).[6,8,9,12,13] Despite this growing evidence, little is known about how the effect of cigarette smoking on CRC survival varies across molecular subtypes of CRC.[8]

We conducted this study to assess if the association between cigarette smoking and CRC survival varies across molecular subtypes of CRC. Additionally, given the growing burden of early-onset CRC (i.e., diagnosed at ages <50 years)[14], and because it has been reported that clinicopathological features, underlying molecular profiles, and risk factors differ between early-onset and later-onset CRC patients[14], we also examined whether the association between smoking and CRC survival differed for early-onset vs. later-onset CRC.

Methods

Study Participants

The present analysis utilized data from two complementary study populations: the Colon Cancer Family Registry (CCFR) and the Advanced Colorectal Cancer of Serrated Subtype (ACCESS) Study. Both study populations have been described in detail elsewhere.[15-17] In summary, the CCFR, which constituted the majority of our study population, is an international collaborative effort including population-based recruitment sites in Australia (Melbourne, Victoria), Canada (Toronto, Ontario), and the United States (Seattle, Washington; Mayo Clinic, Minnesota);[15,16] cases from the CCFR included in the present analysis were diagnosed with incident invasive CRC between 1997-2008 at ages ranging from 18-74 years. The ACCESS Study was a single-site population-based study conducted in western Washington State, including persons diagnosed with incident invasive CRC between 2016-2018 at ages ranging from 18-74 years.[17]

Participants from the CCFR and ACCESS who completed a baseline questionnaire at the time of enrollment, and for whom data on at least one of four tumor markers were available (see below) were included in this analysis. The baseline questionnaire used in the ACCESS Study was

based on the CCFR questionnaire; both questionnaires included questions pertaining to smoking history and, as applicable, smoking patterns from a reference timepoint specified as two years before CRC diagnosis. Baseline questionnaires also included items pertaining to demographic factors (age, sex, height, weight, marital status, education), aspects of past and current medical history (e.g., prior diagnosis of diabetes), physical activity, regular use of aspirin and other non-steroidal anti-inflammatory drug (NSAID), alcohol consumption, and aspects of diet.[15-17]

Cigarette Smoking

Study participants were requested to provide information about their cigarette smoking status two years prior to baseline interview and at the time of their diagnosis. Participants who had consumed at least one cigarette per day for a minimum period of three months were considered as ever smokers; whereas those who had never smoked cigarettes, and those who had smoked at least one cigarette per day but for a period less than three months were considered as never smokers. Smokers were further subdivided into two groups: current smokers and former smokers depending on whether they were actively smoking during the first time they responded to the questionnaires. Current and former smokers were asked to provide further information on the dates they first started and quit smoking, whichever applied to them.

To model for cigarette smoking intensity in this study, we used information on the average number of cigarettes consumed per day and the number of years of smoking for each current and former smoker participant in the study, and we calculated the number of pack-years (none, <20, 20-39, and ≥40) (defined as the average number of cigarettes used per day divided by 20 and multiplied by the number of years smoked).The number of years since cigarette smoking cessation before the diagnosis (i.e., cigarette abstention years) was calculated using the self-reported ages when the concerned participants reported having quit smoking and their corresponding ages at the time of their diagnosis (≤0, 1-10, 11-20, >20, none).

Outcome Information

Details regarding vital status and cause of death data collection for the CCFR are provided elsewhere.[11] For the ACCESS Study, vital status and cause and date of death were determined via linkage to the Puget Sound Surveillance Epidemiology and End Results cancer registry. For both study populations, cases of CRC death referred to those with an underlying cause of death directly attributed to CRC and its complications, and this was ascertained by referring to the International Classification of Disease (ICD-10), tenth revision codes: C18.0-C.20.0 or C26.0. Given the established relationship of cigarette smoking with several other causes of death, we focused our analysis on the primary outcome of death attributed to CRC. Participants who died due to causes other than CRC were considered as censored at the time of their death.

Colorectal Cancer Molecular Subtypes

We grouped participants according to the classification of molecular subtypes adapted from classifications first proposed by Dr. Jeremy Jass.[10,11] This modified classification has five distinct subtypes of CRC based on the combination of four molecular markers: MSI status, CIMP status, somatic *BRAF* gene mutation status, and somatic *KRAS* gene mutation status. Each subtype was characterized by the combined profile of four molecular markers as follows: type 1 (CIMP-high, MSI-high, mutated *BRAF* and *KRAS* wildtype), type 2 (CIMP-high, MSI-low or stable (MSS), mutated *BRAF*, and *KRAS* wildtype), type 3 (CIMP-low or negative, MSI-low or MSS, *BRAF* wildtype and mutated *KRAS*), type 4 (CIMP-low or negative, MSI-low or MSS, *BRAF* and *KRAS* wildtype), type 5 (CIMP-negative, MSI-high, *BRAF* and *KRAS* wildtype).[11]

Further details on how the individual biomarkers that comprise these subtypes were determined have been provided elsewhere. In summary, MSI status was assessed using either the 10-marker Bethesda panel[11] or a 4-marker immunohistochemistry panel for DNA mismatch repair proteins.[18] In both the CCFR and the ACCESS Studies, testing for mutations in *KRAS* was restricted to sequencing of *KRAS* exon 2,[19] and testing for mutations in *BRAF* was restricted to the predominant c.1799T>A (p.V600E) mutation.[20] CIMP status was classified based on a validated five-gene panel quantitative DNA methylation assay for participants from the C-CFR[21,22]; however, for participants in the ACCESS Study, and a small subset of participants from the C-CFR, CIMP status was classified using an expanded eight-marker CIMP panel.[22]

Other covariates

We performed multivariate analyses to examine the association of smoking with CRC-specific survival within different patient strata. Covariates included in our analyses were selected *a priori* as those attributes most likely to confound the association between cigarette smoking and CRC mortality.[11-12,23-24] Specifically, we included the following covariates in our model: sex, regular use of aspirin, body mass index (BMI), year of diagnosis, and study site. BMI was calculated using participant's height at the time of the diagnosis and their estimated weight two years before the diagnosis of CRC and was categorized according to World Health Organization groupings (<18.5 kg/m^2, 18.5-24.9 kg/m^2, 25.0-29.9 kg/m^2, 30+ kg/m^2). CRC anatomic sites were classified as proximal colon, distal colon, and rectal. Considering the changes in cigarette smoking habits as well as the innovations in treatment of CRC, we accounted for years of diagnosis categorically (≤2000, 2001-2010, >2010).

Statistical Analysis

Missing Values

To address the high proportion of missing values for certain variables in the initial database, we performed the multiple imputation by chained equation method (MICE). The number of rounds of multiple imputation was set at 10. To prevent feedback between different versions of the same variables during the imputation process, we set the ridge parameter to 0.01 to make the algorithm more robust. We performed the multiple imputation as well as the rest of the analysis using R software version 2022.120.

Data Analysis

We used the Pearson Chi-square tests of independence to assess whether there was any significant difference in the distribution of cigarette smoking status by age at diagnosis as a categorical variable (≤50 years and > 50 years) and by different molecular subtypes of CRC. To describe CRC survival as associated with smoking status overall and by groups defined by age at diagnosis and molecular subtypes, we used the Kaplan-Meier estimators to plot the survival curves overall and the survival curves for each age and molecular subtype category, and we used univariate and multivariate Cox Proportional Hazards regression analyses.

For the primary analysis, we used the exposure groups defined by cigarette smoking status, and we performed univariate (unadjusted) and multivariate (adjusted) analyses. We first investigated whether the association between cigarette smoking status and CRC survival varied by groups defined by age at diagnosis, adjusting for the different covariates. We assessed this association successively among those with early (< 50 years) and those with later CRC diagnosis (≥50 years). We also tested for a significant interaction between cigarette smoking status and age at diagnosis adjusting for the same variables enumerated above.

To assess if the same association differs by groups defined by CRC molecular subtypes, we used the same variables for adjustment, and we performed the analysis within each subgroup of participants defined by the molecular subtypes. In addition to covariates described above, we also adjusted these subtype-specific models for age at diagnosis as a continuous variable.

For the secondary analyses, we used exposure groups defined by the number of pack-years of cigarette smoking and the number of years since cigarette smoking cessation as main exposures and we adjusted for the same variables as in the primary analysis.

The proportional hazards assumption was assessed using the Schoenfeld residuals method. To address the violation of the proportional hazards' assumptions for the main exposure cigarette smoking status, we recategorized this variable in two levels: Never smokers and Ever smokers. Current and former smokers were grouped as "ever smokers" as compared to "never smokers" and the assumption was not violated.

All regression models used time since CRC diagnosis as the time axis, with delayed entr at the time of study enrollment. Participant follow-up was continued until either death, loss to follow up or the end of the follow-up period, whichever occurred first.

The significance level of 0.05 was used for any hypothesis test, and all performed hypothesis tests were two-sided tests, and we used the software R and the R-studio interface version 2022.120.

Results

Of the 4,901 participants included in this study, 1,378 participants died due to CRC during study follow-up (average follow-up period = 13.6 years). Characteristics of the study population are provided in Table 1, overall and according to cigarette smoking status. Overall, 57% of participants were ever smokers. Compared to never smokers, ever smokers were more likely to be male (57% vs. 43%) and to report regular use of aspirin or NSAIDs (32% vs. 25%). Ever smokers with CRC were less likely to have been diagnosed as early-onset CRC as compared to never smokers with CRC (27% vs. 37%), respectively), and were similar to never smokers with regard to BMI. With respect to tumor characteristics, the distribution of molecular subtype classifications was similar in ever vs. never smokers with CRC, with the exception that the proportion of cases with MSI-high positive tumor subtypes (i.e., types 1 and 5) were slightly higher among ever-smokers (Appendix 1). Participants with type 1 CRC had the highest proportion of ever-smoking history (60.1%), and the lowest proportion of 56% was observed in those with type 3 molecular subtype (Appendix 1).

Cigarette Smoking Status and CRC Survival

During the follow-up period, of the total cases of CRC-related deaths that occurred, 62% were ever smokers and 38 % never smokers. The Kaplan-Meier estimators show that participants who had never smoked cigarettes had higher CRC-specific survival than those who have ever smoked cigarettes (Figure 1). There was a 15-month difference in restricted mean survival times between ever and never smokers, with ever smokers having the shorter mean survival time.

Results from the univariate model show that ever smoker participants had 1.27 (95% CI: 1.13-1.42) times higher hazard of dying from CRC than never smokers. Even after adjusting for different covariates, ever smoker participants continued to have 21% higher risk of dying from CRC compared to never smoker patients (HR=1.21; 95% CI: 1.06-1.38) (Table 2). In analyses

stratified by age at diagnosis, smoking history was not significantly associated with CRC-specific survival among those with early-onset CRC (HR=1.06; 95% CI: 0.81-1.40), whereas a history of ever smoking was associated with significantly poorer survival among those with late-onset CRC (HR=1.26; 95% CI: 1.06-1.48). While notable, these observed differences did not translate to statistically significant interaction by age at diagnosis (p-value: 0.38).

Although numbers were limited in case groups defined by less common molecular subtypes, associations of smoking history with CRC-specific survival were consistent regardless of molecular subtype; however, the observed association was only statistically significant with regard to the predominant CRC type 4 (HR=1.23; 95% CI: 1.03-1.47) (Table 2).

Cigarette Smoking Pack-years, Smoking Cessation and CRC Survival

The effect of smoking cigarettes in terms of number of pack-years of smoking showed that the risk of dying from CRC among ever smokers increased with the number of pack-years. When compared to never smokers, participants with history of ≥40 pack-years of smoking had 52% higher hazard of dying from CRC, followed by participants with history of 20-39 pack-years (24% higher hazards) and then those with <20 pack-years (9% higher hazards of dying from CRC), with a statistically significant positive trend across categories (p-value<0.001) (Table 3). Similar patterns of association were observed among cases with late-onset CRC and those with subtype 4 CRC.

No statistically significant association between smoking and CRC survival was observed among those with early-onset CRC, although there were suggestive non-significant associations in comparisons of non-smokers with those in the highest pack-years groups.

Even though the associations between time since smoking cessation and CRC survival were not statistically significant, regardless of age at diagnosis and molecular subtype groups, there was a statistically significant decreasing trend of CRC deaths across levels of time since smoking cessation variable (p-value<0.001) (Table 5 & Table 6).

Discussion

We explored whether pre-diagnostic smoking history affects CRC-specific survival, and whether this association differs between early-onset and late-onset CRC patients and according to CRC molecular subtype groups. Our findings of poorer CRC-specific survival associated with history of smoking are consistent with the results found by Phipps et al, who have reported a statistically significant difference in hazard of CRC-specific deaths between ever-smokers and never smokers, with ever smokers having 21% higher hazard[6]. These results also corroborate the results previously reported by Munro, A.J et al who have found 21% higher hazard of CRC-specific death when comparing active smokers to non-smokers[25.] Several other studies have found associations of similar magnitude.[23,26] Our findings also corroborate the results reported by others indicating an increase in the hazards of death from CRC with increasing cumulative pack-years of smoking.[6,26 27]

Age-stratified results from our study are identical to the results reported by Phipps, A. I et al[6] whose study sample is a subset of the current study (N=2,933)**.** We found statistically significant difference in hazards of dying from CRC between ever and never smokers only among patients with late-onset CRC[6]. For the stratified analysis by molecular subtype groups, to our knowledge, our study is the first one to assess the effect of smoking cigarette on CRC survival by molecular subtypes considering the four biomarkers together. However, despite the absence of statistically significant difference in hazards of death from CRC observed between ever smokers

and never smokers among participants with different molecular subtypes except for those with subtype 4, our results show that the association with survival was mostly similar in magnitude across the different molecular subtypes. This absence of statistically significant results could be explained by the small number of participants within these molecular subtype groups.

Previous studies have shown tremendous evidence that cigarette smoking was associated with these CRC biomarkers (MSI status, CIMP status, and BRAF mutation status).[28-29] Given that these tumor attributes considered individually or combined in subtypes are also known to be associated with survival among patients with CRC, it is plausible that cigarette smoking could influence CRC survival through an impact on tumor biology.[11,19,20,30]

Results from our study show that globally, the risk of death from CRC increases with the number of pack-years of smoking when comparing ever-smokers to never-smokers.

Strengths and Limitations

These findings should be interpreted in the context of study limitations. In particular, although efforts were made to enroll participants within a relatively short timeframe after diagnosis, there was a time lapse of several months; as such, a number of otherwise eligible individuals likely died before they could enroll in the CCFR or ACCESS studies. Survivor bias is possible to the extent that such individuals differed from included participants with respect to their smoking history. Cigarette smoking status, the average number of cigarettes consumed daily, and the duration of smoking were all self-reported, which may have led to recall bias. Finally, not fully accounting for key factors known to be linked with patient survival, such as socioeconomic level and other comorbidities is a drawback of this study.

Despite the aforementioned limitations, our study has a number of noteworthy advantages to consider. The study was conducted with a population-based sample of 4,901 participants from three countries. The generalization of this study's findings could be aided by incorporating data from multiple study centers in different nations. According to our knowledge, this study is the first to examine the effects of smoking on survival by molecular subtypes accounting for the four most commonly used molecular biomarkers of CRC tumors. The availability of data on potential confounders and demographic variables has made it possible to derive less confounded estimates of survival for each exposure group by adjusting for their effects.

Because smoking cigarettes is associated with both the risk of developing colorectal cancer and the higher hazards of death from CRC, considerable efforts need to be deployed to promote smoking cessation among people who have a higher risk of developing CRC as well as among those who have already been diagnosed with CRC, regardless of their age at diagnosis and the molecular subtypes, in order to improve the overall CRC-specific survival.

Tables and Figures

Table 1. Demographic and Clinicopathologic Characteristics of Study Participants by Smoking Status

	Never Smoker	Ever Smoker	Overall
Sex			
Male	43.3%	57.4%	51.4%
Female	56.7%	42.6%	48.6%
Age at Diagnosis			
Mean (SD)	54.6 (11.9)	57.3 (11.0)	56.1 (11.5)
<50yrs	36.8%	27.4%	31.4%
≥50yrs	63.2%	72.6%	68.6%
Aspirin Use			
Yes	25.0%	31.9%	28.9%
No	75.0%	68.1%	71.1%
BMI			
Mean (SD)	27.6 (6.00)	27.7 (5.61)	27.6 (5.78)
Under weight	1.7%	1.8%	1.7%
Normal weight	36.3%	30.9%	33.2%
Overweight	35.6%	40.3%	38.3%
Obesity	26.4%	27.0%	26.7%
CRC sites			
Proximal	37.2%	34.5%	35.7%
Distal	30.6%	30.6%	30.6%
Rectal	32.2%	34.9%	33.7%
Molecular Subtypes			
type1	5.9%	6.8%	6.4%
type2	2.9%	2.9%	2.9%
type3	31.5%	29.8%	30.5%
type4	53.5%	54.6%	54.1%
type5	6.2%	6.0%	6.1%
Diagnosis Years			
≤ 2000	52.9%	62.9%	58.6%
> 2000	47.1%	37.1%	41.4%
CRC Deaths			
alive	75.1%	69.5%	71.9%
dead	24.9%	30.5%	28.1%

Note: Only percentages are used because of multiple imputation

Table 2. Smoking Status and CRC Survival by Age at Diagnosis and by Molecular Subtypes

	Unadjusted Models	**Adjusted Models (*)**
Overall	NS: Ref. ES: HR= 1.27(1.13-1.42) <0.001	NS: Ref. ES:HR= 1.21(1.06-1.38) 0.0046
Age Groups		
Early-Onset	NS: Ref. ES: HR= 1.16(0.95-1.41) 0.14	NS: Ref. ES: HR= 1.06(0.80-1.40) 0.66
Late-Onset	NS: Ref. ES: HR= 1.32(1.14-1.52) <0.001	NS: Ref. ES: HR= 1.26(1.06-1.50) 0.007
Molecular Subtypes		
Type1	NS Ref. ES: HR=1.57(0.76-3.22) 0.21	NS: Ref. ES: HR=1.67(0.69-4.03) 0.24
Type2	NS: Ref. ES: HR=1.39(0.62-3.14) 0.40	NS: Ref. ES: HR=0.70(0.19-2.55) 0.57
Type3	NS: Ref. ES: HR=1.25(1.01-1.56) 0.042	NS: Ref. ES: HR=1.18(0.92-1.52) 0.19
Type4	NS: Ref. ES: HR=1.28(1.09-1.51) 0.002	NS: Ref. ES: HR=1.23(1.03-1.47) 0.025
Type5	NS: Ref. ES: HR=1.07(0.56-2.06) 0.82	NS: Ref. ES: HR=1.17(0.46-2.99) 0.73

Note: NS= Never smoker ES= Ever smoker
Type 1 (CIMP-high, MSI-high, mutated BRAF and KRAS wildtype)
Type 2 (CIMP-high, MSI-low or stable (MSS), mutated BRAF, and KRAS wildtype)
Type 3 (CIMP-low or negative, MSI-low or MSS, BRAF wildtype and mutated KRAS)
Type 4 (CIMP-low or negative, MSI-low or MSS, BRAF and KRAS wildtype)
Type 5 (CIMP-negative, MSI-high, BRAF and KRAS wildtype)
(*) We adjusted for sex, age at diagnosis, aspirin use, body mass index, CRC sites, molecular subtypes, and diagnosis years.

Table 3. Smoking Pack-years and CRC Survival by Age Groups at Diagnosis

Pack-Years	**Adjusted Models (*)**	
Overall		
Never-Smokers	Ref.	
<20 Pack-Years	HR= 1.09(0.94-1.28)	0.25
20-39 Pack-Years	HR= 1.24(1.03-1.50)	0.02
≥40 Pack-Years	HR= 1.52(1.24-1.87)	<0.001
p-value for trend test <0.001		
Early-Onset		
Never-Smokers	Ref.	
<20 Pack-Years	HR=0.94(0.69-1.28)	0.68
20-39 Pack-Years	HR= 1.40(0.90-2.18)	0.13
≥40 Pack-Years	HR= 1.39(0.63-3.09)	0.42
Late-Onset		
Never-Smokers	Ref.	
<20 Pack-Years	HR=1.14(0.93-1.39)	0.20
20-39 Pack-Years	HR=1.22(0.98-1.52)	0.079
≥40 Pack-Years	HR=1.52(1.21-1.92)	0.037

Note: Ref. indicates the reference group
(*) We adjusted for sex, age at diagnosis, aspirin use, body mass index, CRC sites, molecular subtypes, and diagnosis years.

Table 4. Smoking Pack-years and CRC Survival by Molecular Subtypes

PACK-YEARS	ADJUSTED MODELS (*)	
Type1		
Never-Smokers	Ref.	
<20 Pack-Years	HR= 1.43(0.41-5.03)	0.55
20-39 Pack-Years	HR= 1.16(0.34-3.92)	0.81
≥40 Pack-Years	HR= 2.83(0.95-8.43)	0.06
Type2		
Never-Smokers	Ref	
<20 Pack-Years	HR= 0.72(0.15-3.46)	0.67
20-39 Pack-Years	HR= 0.62(0.07-5.76)	0.65
≥40 Pack-Years	HR= 0.57(0.06-5.80)	0.62
Type3		
Never-Smokers	Ref.	
<20 Pack-Years	HR= 1.04(0.76-1.44)	0.79
20-39 Pack-Years	HR= 1.21(0.88-1.67)	0.24
≥40 Pack-Years	HR=1.60(0.99-2.55)	0.05
Type4		
Never-Smokers	Ref.	
<20 Pack-Years	HR=1.11(0.89-1.38)	0.37
20-39 Pack-Years	HR=1.30(1.02-1.66)	0.031
≥40 Pack-Years	HR= 1.46(1.11-1.92)	0.006
Type5		
Never-Smokers	Ref.	
<20 Pack-Years	HR= 1.27(0.42-3.81)	0.66
20-39 Pack-Years	HR=0.87(0.21-3.53)	0.84
≥40 Pack-Years	HR=1.18(0.14-9.78)	0.86

Note: Ref. indicates the reference group
Type 1 (CIMP-high, MSI-high, mutated BRAF and KRAS wildtype)
Type 2 (CIMP-high, MSI-low or stable (MSS), mutated BRAF, and KRAS wildtype)
Type 3 (CIMP-low or negative, MSI-low or MSS, BRAF wildtype and mutated KRAS)
Type 4 (CIMP-low or negative, MSI-low or MSS, BRAF and KRAS wildtype)
Type 5 (CIMP-negative, MSI-high, BRAF and KRAS wildtype)
(*) We adjusted for sex, age at diagnosis, aspirin use, body mass index, CRC sites, molecular subtypes, and diagnosis years.

Table 5. Smoking Cessation Years and CRC Survival by Early and Late-onset Status

	Adjusted Models (*)	
Overall		
Continued to Smoke	(Ref)	
Quit < 10 y	HR= 0.86(0.64-1.15)	0.31
Quit 10-20 y	HR= 0.86(0.65-1.15)	0.30
Quit > 20 y	HR= 0.80(0.60-1.07)	0.13
Never Smoked	HR=0.71(0.56-0.92)	0.008
p-value trend test <0.001		
< 50 Years		
Continued to Smoke	(Ref)	
Quit < 10 y	HR=0.84(0.48-1.48)	0.55
Quit 10-20 y	HR=1.14(0.61-2.12)	0.67
Quit > 20 y	HR=1.05(0.50-2.19)	0.89
Never Smoked	HR=0.95(0.55-1.61)	0.83
≥ 50 Years		
Continued to Smoke	(Ref)	
Quit < 10 y	HR=0.88(0.62-1.26)	0.48
Quit 10-20 y	HR=0.77(0.53-1.14)	0.18
Quit > 20 y	HR=0.72(0.51-1.01)	0.05
Never Smoked	HR=0.64(0.47-0.88)	0.005

Note: Ref. indicates the reference group
(*) We adjusted for sex, age at diagnosis, aspirin use, body mass index, CRC sites, molecular subtypes, and diagnosis years.

Table 6. Smoking Cessation Years and Survival by Molecular Subtypes*

	Adjusted Models (*)	
Type1		
Continued To Smoke	Ref.	
Quit < 10 Y	HR=0.28(0.05-1.69)	0.15
Quit 10-20 Y	HR=1.14(0.61-2.12)	0.67
Quit > 20 Y	HR=1.05(0.50-2.19)	0.89
Never Smoked	HR=0.95(0.55-1.61)	0.83
Type3		
Continued To Smoke	Ref.	
Quit < 10 Y	HR=0.90(0.50-1.64)	0.72
Quit 10-20 Y	HR=0.80(0.50-1.27)	0.34
Quit > 20 Y	HR=0.82(0.49-1.38)	0.45
Never Smoked	HR=0.72(0.47-1.13)	0.15
Type4		
Continued To Smoke	Ref.	
Quit < 10 Y	HR=0.94(0.61-1.44)	0.76
Quit 10-20 Y	HR=1.01(0.67-1.52)	0.97
Quit > 20 Y	HR=0.81(0.54-1.21)	0.29
Never Smoked	HR=0.75(0.53-1.07)	0.11
Type5		
Continued To Smoke	Ref.	
Quit < 10 Y	HR= 0.78(0.09-6.08)	0.80
Quit 10-20 Y	HR=0.54(0.03-9.00)	0.63
Quit > 20 Y	HR=1.53(0.28-8.50)	0.61
Never Smoked	HR=0.84(0.17-4.06)	0.82

Ref. indicates the reference group.
*Small numbers precluded analysis for Type 2 tumors.
Type 1 (CIMP-high, MSI-high, mutated BRAF and KRAS wildtype)
Type 3 (CIMP-low or negative, MSI-low or MSS, BRAF wildtype and mutated KRAS)
Type 4 (CIMP-low or negative, MSI-low or MSS, BRAF and KRAS wildtype)
Type 5 (CIMP-negative, MSI-high, BRAF and KRAS wildtype)
(*) We adjusted for sex, age at diagnosis, aspirin use, body mass index, CRC sites, molecular subtypes, and diagnosis years.

Appendix1. Distribution of Participants Smoking Status by Molecular Subtypes

	type1	type2	type3	type4	type5	Overall
Smoking Status						
Never Smoker	39.5%	43.0%	44.1%	42.2%	43.4%	42.7%
Ever Smoker	60.5%	57.0%	55.9%	57.8%	56.6%	57.3%

Note: Only percentages are used because of multiple imputation

Appendix2. Distribution of Participants Smoking Status by Age at Diagnosis

	<50yrs	≥50yrs	Overall
Smoking Status			
Never Smoker	50.0%	39.4%	42.7%
Ever Smoker	50.0%	60.6%	57.3%

Note: Only percentages are used because of multiple imputation

Appendix3. Distribution of Participants Molecular Subtypes by Age at Diagnosis Groups

	<50yrs	≥50yrs	Overall
Molecular subtypes			
type1	0.9%	8.9%	6.4%
type2	2.2%	3.2%	2.9%
type3	30.7%	30.4%	30.5%
type4	56.3%	53.1%	54.1%
type5	9.9%	4.3%	6.1%

Note: Only percentages are used because of multiple imputation

Figure1.

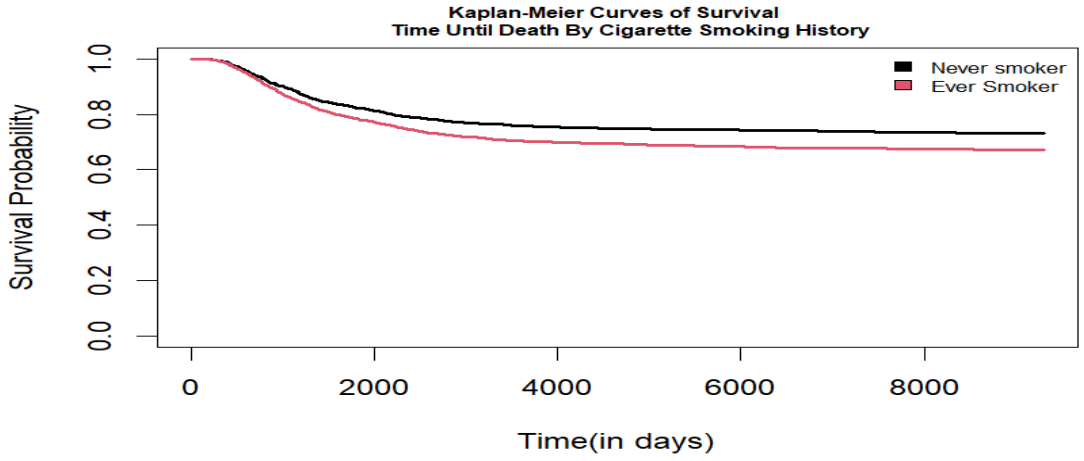


Figure 2.

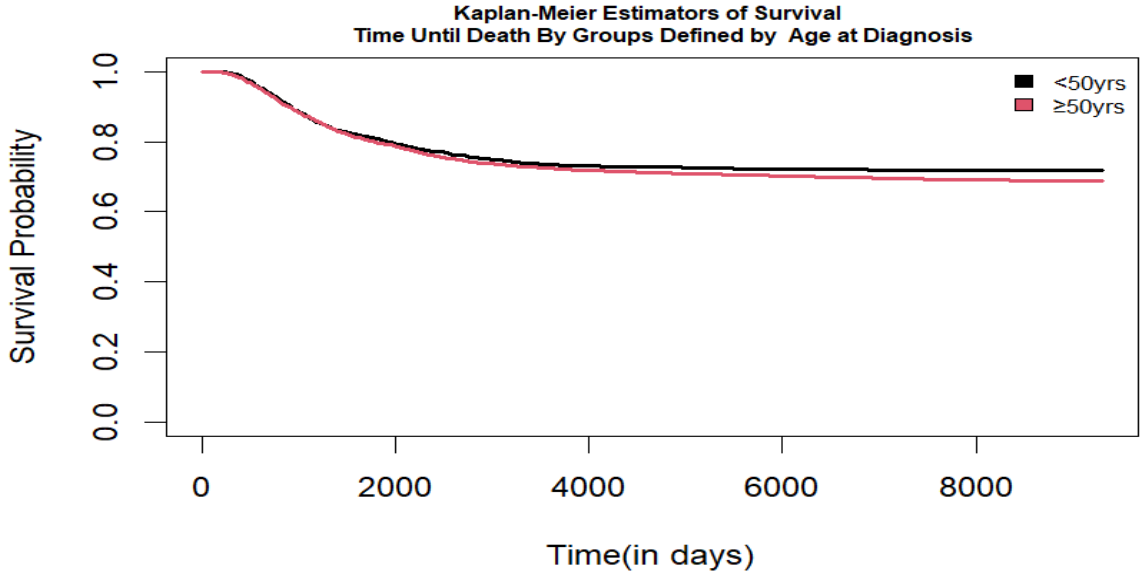
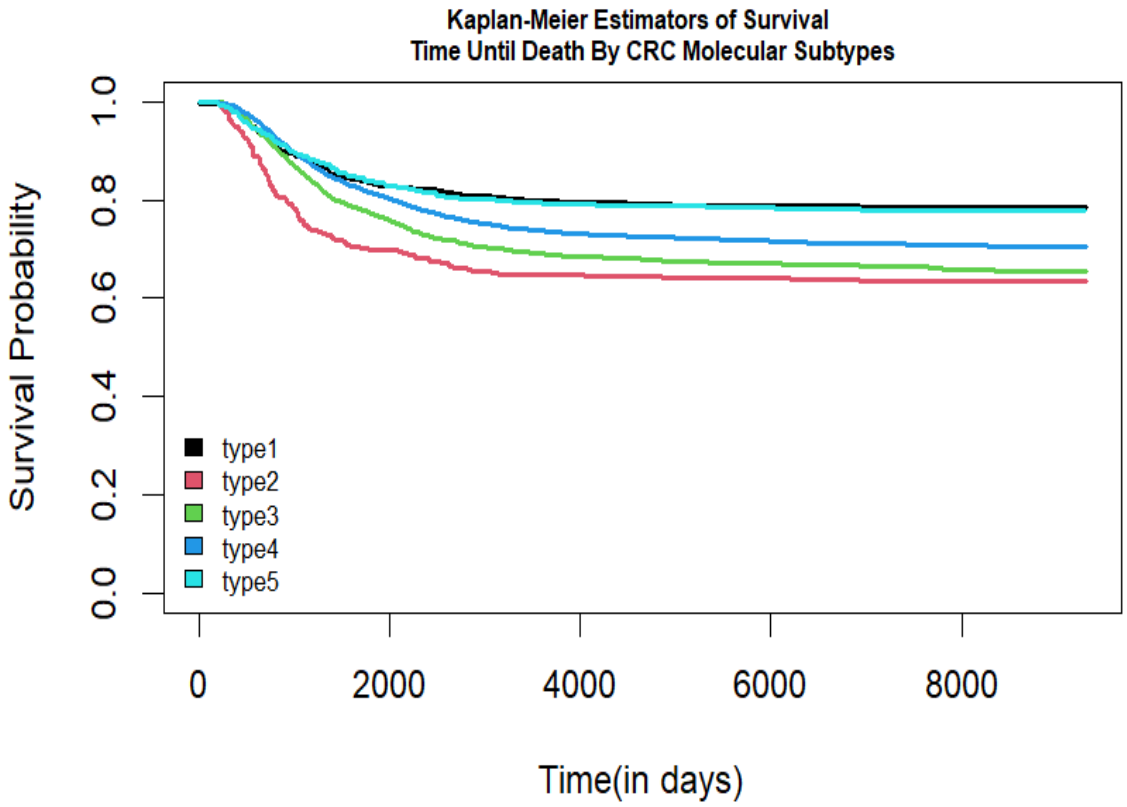


Figure 3.


Kaplan-Meier Estimators of Survival
Time Until Death By CRC Molecular Subtypes
Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
type1
type2
type3
type4
type5
0
2000
4000
6000
8000
Time(in days)

References

1. USCS Data Visualizations - CDC

2. Colorectal Cancer Survival Rates | Colorectal Cancer Prognosis

3. Li, N., Lu, B., Luo, C., Cai, J., Lu, M., Zhang, Y., Chen, H., & Dai, M. (2021). Incidence, mortality, survival, risk factor and screening of colorectal cancer: A comparison among China, Europe, and northern America. Cancer letters, 522, 255–268. https://doi.org/10.1016/j.canlet.2021.09.034

4.Leong, E., Ong, S. K., Madli, F., Tan, A., Lai, D., Basir, N., Ramlee, N., & Chong, V. H. (2020). Survival Rates and Associated Factors of Colorectal Cancer Patients in Brunei Darussalam. Asian Pacific journal of cancer prevention: APJCP, 21(1), 259–265. https://doi.org/10.31557/APJCP.2020.21.1.259

5.Coughlin S. S. (2020). Social determinants of colorectal cancer risk, stage, and survival: a systematic review. International journal of colorectal disease, 35(6), 985–995. https://doi.org/10.1007/s00384-020-03585-z

6.Phipps, A. I., Baron, J., & Newcomb, P. A. (2011). Prediagnostic smoking history, alcohol consumption, and colorectal cancer survival: the Seattle Colon Cancer Family Registry. Cancer, 117(21), 4948–4957. https://doi.org/10.1002/cncr.26114

7.Huang, Y. M., Wei, P. L., Ho, C. H., & Yeh, C. C. (2022). Cigarette Smoking Associated with Colorectal Cancer Survival: A Nationwide, Population-Based Cohort Study. Journal of clinical medicine, 11(4), 913. https://doi.org/10.3390/jcm11040913

8.Zhu, Y., Yang, S. R., Wang, P. P., Savas, S., Wish, T., Zhao, J., Green, R., Woods, M., Sun, Z., Roebothan, B., Squires, J., Buehler, S., Dicks, E., Zhao, J., Mclaughlin, J. R., Parfrey, P. S., & Campbell, P. T. (2014). Influence of pre-diagnostic cigarette smoking on colorectal cancer survival: overall and by tumour molecular phenotype. British journal of cancer, 110(5), 1359–1366. https://doi.org/10.1038/bjc.2014.6

9.Botteri, E., Borroni, E., Sloan, E. K., Bagnardi, V., Bosetti, C., Peveri, G., Santucci, C., Specchia, C., van den Brandt, P., Gallus, S., & Lugo, A. (2020). Smoking and Colorectal Cancer Risk, Overall and by Molecular Subtypes: A Meta-Analysis. The American journal of gastroenterology, 115(12), 1940–1949. https://doi.org/10.14309/ajg.0000000000000803

10.Jass J. R. (2007). Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. Histopathology, 50(1), 113–130. https://doi.org/10.1111/j.1365-2559.2006.02549.

11.Phipps, A. I., Limburg, P. J., Baron, J. A., Burnett-Hartman, A. N., Wiesenberger, D. J., Laird, P. W., Sinicrope, F. A., Rosty, C., Buchanan, D. D., Potter, J. D., & Newcomb, P. A. (2015). Association between molecular subtypes of colorectal cancer and patient survival. Gastroenterology, 148(1), 77–87. e2. https://doi.org/10.1053/j.gastro.2014.09.038

12.Wang, X., Amitay, E., Harrison, T. A., Banbury, B. L., Berndt, S. I., Brenner, H., Buchanan, D. D., Campbell, P. T., Cao, Y., Chan, A. T., Chang-Claude, J., Gallinger, S. J., Giannakis, M., Giles, G. G., Gunter, M. J., Hopper, J. L., Jenkins, M. A., Lin, Y., Moreno, V., Nishihara, R., … Peters, U. (2021). Association Between Smoking and Molecular Subtypes of Colorectal Cancer. JNCI cancer spectrum, 5(4), pkab056. https://doi.org/10.1093/jncics/pkab056

13.Minami, Y., Kanemura, S., Kusaka, J., Kinouchi, M., Suzuki, S., Nishino, Y., & Miura, K. (2022). Associations of cigarette smoking, alcohol drinking and body mass index with survival after colorectal cancer diagnosis by anatomic subsite: a prospective patient cohort study in Japan. Japanese journal of clinical oncology, 52(12), 1375–1388. https://doi.org/10.1093/jjco/hyac140

14.REACCT Collaborative, Zaborowski, A. M., Abdile, A., Adamina, M., Aigner, F., d'Allens, L., Allmer, C., Álvarez, A., Anula, R., Andric, M., Atallah, S., Bach, S., Bala, M., Barussaud, M., Bausys, A., Bebington, B., Beggs, A., Bellolio, F., Bennett, M. R., Berdinskikh, A., … Winter, D. C. (2021). Characteristics of Early-Onset vs Late-Onset Colorectal Cancer: A Review. JAMA surgery, 156(9), 865–874. https://doi.org/10.1001/jamasurg.2021.2380

15.Newcomb, P. A., Baron, J., Cotterchio, M., Gallinger, S., Grove, J., Haile, R., Hall, D., Hopper, J. L., Jass, J., Le Marchand, L., Limburg, P., Lindor, N., Potter, J. D., Templeton, A. S., Thibodeau, S., Seminara, D., & Colon Cancer Family Registry (2007). Colon Cancer Family Registry: an international resource for studies of the genetic epidemiology of colon cancer. Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 16(11), 2331–2343. https://doi.org/10.1158/1055-9965.EPI-07-0648

16.Phipps, A. I., Robinson, J. R., Campbell, P. T., Win, A. K., Figueiredo, J. C., Lindor, N. M., & Newcomb, P. A. (2017). Prediagnostic alcohol consumption and colorectal cancer survival The Colon Cancer Family Registry. Cancer, 123(6), 1035–1043. https://doi.org/10.1002/cncr.30446

17.https://www.fredhutch.org/en/research/divisions/public-health-sciences-division/research/cancer-prevention/

18.Shia J. (2008). Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. The Journal of molecular diagnostics : JMD, 10(4), 293–300. https://doi-org.offcampus.lib.washington.edu/10.2353/jmoldx.2008.080031

19.Phipps, A. I., Buchanan, D. D., Makar, K. W., Win, A. K., Baron, J. A., Lindor, N. M., Potter, J. D., & Newcomb, P. A. (2013). KRAS-mutation status in relation to colorectal cancer survival: the joint impact of correlated tumour markers. British journal of cancer, 108(8), 1757–1764. https://doi-org.offcampus.lib.washington.edu/10.1038/bjc.2013.118

20.Phipps, A. I., Buchanan, D. D., Makar, K. W., Burnett-Hartman, A. N., Coghill, A. E., Passarelli, M. N., Baron, J. A., Ahnen, D. J., Win, A. K., Potter, J. D., & Newcomb, P. A. (2012). BRAF mutation status and survival after colorectal cancer diagnosis according to patient and tumor characteristics. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 21(10), 1792–1798. https://doi-org.offcampus.lib.washington.edu/10.1158/1055-9965.EPI-12-0674

21. Levine, A. J., Phipps, A. I., Baron, J. A., Buchanan, D. D., Ahnen, D. J., Cohen, S. A., Lindor, N. M., Newcomb, P. A., Rosty, C., Haile, R. W., Laird, P. W., & Weisenberger, D. J. (2016). Clinicopathologic Risk Factor Distributions for MLH1 Promoter Region Methylation in CIMP-Positive Tumors. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 25(1), 68–75. https://doi-org.offcampus.lib.washington.edu/10.1158/1055-9965.EPI-15-0935

22.Ogino, S., Kawasaki, T., Kirkner, G. J., Kraft, P., Loda, M., & Fuchs, C. S. (2007). Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. The Journal of molecular diagnostics : JMD, 9(3), 305–314. https://doi.org/10.2353/jmoldx.2007.060170

23.Jayasekara, H., English, D. R., Haydon, A., Hodge, A. M., Lynch, B. M., Rosty, C., Williamson, E. J., Clendenning, M., Southey, M. C., Jenkins, M. A., Room, R., Hopper, J. L., Milne, R. L., Buchanan, D. D., Giles, G. G., & MacInnis, R. J. (2018). Associations of alcohol intake, smoking, physical activity and obesity with survival following colorectal cancer diagnosis by stage, anatomic site and tumor molecular subtype. International journal of cancer, 142(2), 238–250. https://doi.org/10.1002/ijc.31049

24.Samowitz, W. S., Curtin, K., Schaffer, D., Robertson, M., Leppert, M., & Slattery, M. L. (2000). Relationship of Ki-ras mutations in colon cancers to tumor location, stage, and survival: a population-based study. Cancer epidemiology, biomarkers & prevention 9: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 9(11), 1193–1197.

25.Munro, A. J., Bentley, A. H., Ackland, C., & Boyle, P. J. (2006). Smoking compromises cause-specific survival in patients with operable colorectal cancer. Clinical oncology (Royal College of Radiologists (Great Britain)), 18(6), 436–440. https://doi.org/10.1016/j.clon.2006.04.009

26. Jang, D., Choe, S., Park, J. W., Jeong, S. Y., & Shin, A. (2020). Smoking status before and after colorectal cancer diagnosis and mortality in Korean men: A population-based cohort study. *Cancer medicine*, *9*(24), 9641–9648. https://doi.org/10.1002/cam4.3609

27.Parajuli, R., Bjerkaas, E., Tverdal, A., Le Marchand, L., Weiderpass, E., & Gram, I. T. (2014). Cigarette smoking and colorectal cancer mortality among 602,242 Norwegian males and females. Clinical epidemiology, 6, 137–145. https://doi.org/10.2147/CLEP.S58722

28. Botteri, E., Borroni, E., Sloan, E. K., Bagnardi, V., Bosetti, C., Peveri, G., Santucci, C., Specchia, C., van den Brandt, P., Gallus, S., & Lugo, A. (2020). Smoking and Colorectal Cancer Risk, Overall and by Molecular Subtypes: A Meta-Analysis. The American journal of gastroenterology, 115(12), 1940–1949. https://doi.org/10.14309/ajg.0000000000000803

29. Poynter, J. N., Haile, R. W., Siegmund, K. D., Campbell, P. T., Figueiredo, J. C., Limburg, P., Young, J., Le Marchand, L., Potter, J. D., Cotterchio, M., Casey, G., Hopper, J. L., Jenkins, M. A., Thibodeau, S. N., Newcomb, P. A., Baron, J. A., & Colon Cancer Family Registry (2009). Associations between smoking, alcohol consumption, and colorectal cancer, overall and by tumor microsatellite instability status. Cancer epidemiology, biomarkers & prevention: a publication of

the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology, 18(10), 2745–2750. https://doi.org/10.1158/1055-9965.EPI-09-0517

30. Phipps, A. I., Alwers, E., Harrison, T., Banbury, B., Brenner, H., Campbell, P. T., Chang-Claude, J., Buchanan, D., Chan, A. T., Farris, A. B., Figueiredo, J. C., Gallinger, S., Giles, G. G., Jenkins, M., Milne, R. L., Newcomb, P. A., Slattery, M. L., Song, M., Ogino, S., Zaidi, S. H., … Peters, U. (2020). Association Between Molecular Subtypes of Colorectal Tumors and Patient Survival, Based on Pooled Analysis of 7 International Studies. *Gastroenterology*, *158*(8), 2158–2168.e4. https://doi.org/10.1053/j.gastro.2020.02.029