



**ARTICLE**

Clinical Study

# Smoking, alcohol consumption and colorectal cancer risk by molecular pathological subtypes and pathways

Efrat L. Amitay<sup>1</sup>, Prudence R. Carr<sup>1</sup>, Lina Jansen<sup>1</sup>, Wilfried Roth<sup>2,3</sup>, Elizabeth Alwers<sup>1,4</sup>, Esther Herpel<sup>3,5</sup>, Matthias Kloor<sup>6</sup>, Hendrik Bläker<sup>7</sup>, Jenny Chang-Claude<sup>8</sup>, Hermann Brenner<sup>1,9,10</sup> and Michael Hoffmeister<sup>1</sup>

**BACKGROUND:** Smoking and alcohol increase risk for colorectal malignancies. However, colorectal cancer (CRC) is a heterogenic disease and associations with the molecular pathological pathways are unclear.

**METHODS:** This population-based case-control study includes 2444 cases with first-diagnosis CRC and 2475 controls. Tumour tissue was analysed for MSI (microsatellite instability), CIMP (CpG island methylator phenotype), BRAF (B-Raf proto-oncogene serine/threonine kinase gene) and KRAS (Kirsten rat sarcoma viral oncogene homologue gene) mutations. Odds ratios (ORs) and 95% confidence intervals (95% CIs) were estimated for associations between alcohol and smoking and CRC molecular subtypes and pathways.

**RESULTS:** Current smoking showed higher ORs for MSI-high (OR = 2.79, 95% CI: 1.86–4.18) compared to MSS (OR = 1.41, 1.14–1.75, *p*-heterogeneity (*p*-het) = 0.001), BRAF-mutated (mut) (OR = 2.40, 1.41–4.07) compared to BRAF-wild type (wt) (OR = 1.52, 1.24–1.88, *p*-het = 0.074), KRAS-wt (OR = 1.70, 1.36–2.13) compared to KRAS-mut (OR = 1.26, 0.95–1.68, *p*-het = 0.039) and CIMP-high (OR = 2.01, 1.40–2.88) compared to CIMP-low/negative CRC (OR = 1.50, 1.22–1.85, *p*-het=0.101). Current smoking seemed more strongly associated with sessile serrated pathway (CIMP-high + BRAF-mut; OR = 2.39, 1.27–4.52) than with traditional pathway CRC (MSS + CIMP-low/negative + BRAF-wt; OR = 1.50, 1.16–1.94) and no association was observed with alternate pathway CRC (MSS + CIMP-low/negative + KRAS-wt; OR = 1.08, 0.77–1.43). No heterogeneity was observed in alcohol consumption association by molecular subtypes.

**CONCLUSIONS:** In this large case-control study, smoking was more strongly associated with MSI-high and KRAS-wt CRC and with cases showing features of the sessile serrated pathway. Association patterns were less clear for alcohol consumption.

*British Journal of Cancer* (2020) 122:1604–1610; <https://doi.org/10.1038/s41416-020-0803-0>

**BACKGROUND**

Often considered one disease, sporadic colorectal cancer (CRC), accounting for 95% of CRC cases, is a heterogeneous disease arising from different sets of genetic and epigenetic alterations.<sup>1</sup> The most established underlying molecular pathological subtypes of CRC are characterised by microsatellite instability (MSI) (prevalence 15% in sporadic CRC), CpG island methylator phenotype (CIMP-high, 20%),<sup>2</sup> B-Raf proto-oncogene serine/threonine kinase gene mutations (BRAF mutations, 10%) and Kirsten rat sarcoma viral oncogene homologue gene mutations (KRAS mutations, 30–50% of sporadic CRC cases). KRAS and BRAF mutations are considered mutually exclusive.<sup>3</sup> Another important molecular feature is the adenomatous polyposis coli (*APC*) gene, a key tumour suppressor gene mutated in 45% to 81% of sporadic CRC cases.<sup>2</sup>

Smoking is associated with increased risk of CRC.<sup>4</sup> Moderate to high intake of alcohol was shown to increase CRC risk in a linear dose-response association.<sup>5</sup> Several studies found smoking was associated with a higher risk of MSI-high,<sup>6–10</sup> CIMP-high,<sup>8,10,11</sup> BRAF-mutated (mut)<sup>8,10,11</sup> and KRAS-wild-type (wt) CRC<sup>12,13</sup> while

others found no differential association by CRC subtypes.<sup>13–16</sup> In one study, alcohol was found<sup>17</sup> to be associated with increased risk of MSI-high compared to MSS CRC. In other previous studies, alcohol was not differentially associated with either MSI,<sup>9,18</sup> CIMP, BRAF<sup>19–21</sup> or KRAS status.<sup>22</sup>

Since smoking and high alcohol consumption are often correlated, both risk factors were investigated in this study. The aim of this study was to extend current knowledge on the associations of smoking and alcohol consumption with major molecular subtypes and pathways of CRC.

**METHODS**

Study population

The DACHS study (Darmkrebs: Chancen der Verhütung durch Screening; CRC: chances for prevention through screening), an ongoing case-control study with follow-up of CRC cases, was initiated in 2003 and has been described in detail previously.<sup>23,24</sup> In short, cases with a first, histologically confirmed, diagnosis of

<sup>1</sup>Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany; <sup>2</sup>Institute of Pathology, University Medical Center Mainz, Mainz, Germany; <sup>3</sup>Institute of Pathology, University Hospital Heidelberg, Heidelberg, Germany; <sup>4</sup>Medical Faculty, Heidelberg University, Heidelberg, Germany; <sup>5</sup>NCT Tissue Bank, National Center for Tumor Diseases (NCT), Heidelberg, Germany; <sup>6</sup>Department of Applied Tumor Biology, Institute of Pathology, University Hospital Heidelberg, Heidelberg, Germany; <sup>7</sup>Institute of Pathology, University hospital Leipzig, Leipzig, Germany; <sup>8</sup>Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany; <sup>9</sup>Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany and <sup>10</sup>German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany  
Correspondence: Efrat L. Amitay (e.amitay@dkfz.de)

Received: 28 November 2019 Revised: 24 February 2020 Accepted: 4 March 2020  
Published online: 30 March 2020

CRC and randomly selected control participants with no history of CRC, frequency matched to cases by age, sex and county of residence, are recruited in the Rhine-Neckar-Odenwald region in Germany (~2 million inhabitants). The DACHS study was approved by the ethics committees of the Medical Faculty of Heidelberg University and the state medical boards of Baden-Wuerttemberg and Rhineland-Palatinate.

#### Data collection

Eligible cases were identified in hospitals within the study region and after giving written informed consent were interviewed by trained interviewers using a standardised questionnaire during hospitalisation after surgery or at home after discharge. The median time between diagnosis and interview was 24 days (interquartile range: 10–224 days). Control participants were randomly selected from population registries and contacted through the study centre to schedule home interviews. Controls with a history of CRC were excluded. Controls opting out of the interview were offered a self-administered short questionnaire. Based on hospital data, ~50% of eligible patients were recruited. The participation rate of eligible controls was 51%.

The current analysis is based on DACHS participants recruited in 2003–2010, as comprehensive molecular tumour analyses of MSI, CIMP, BRAF and KRAS were performed in full for cases joining the study in that period. Participants reporting having Crohn's disease or ulcerative colitis ( $N=28$ ) were excluded from the current analysis. Control participants who answered the short questionnaire only ( $N=658$ ) were also excluded due to missing required information on lifetime alcohol consumption and time of smoking cessation (Supplementary Fig. 1).

#### Assessment of smoking and alcohol consumption

Participants were interviewed regarding smoking history prior to diagnosis (for cases) or interview (controls). Participants were classified as non-smokers if they had never smoked regularly or as former smokers if they had stopped smoking at least 2 years before diagnosis (cases) or interview (controls).

Participants were asked about alcohol consumption in each decade of life from the age of 20 years until diagnosis (cases) or interview (controls). Alcohol consumption was calculated in units of gram ethanol per day. Data were collected on participants' drinking habits of portions of beer (0.33 l), wine (0.25 l) and liquor (0.02 l). Ethanol content was derived from food composition tables,<sup>25</sup> assuming an average of 4, 8.6 and 33 g of pure ethanol in 100 ml of beer, wine and liquor, respectively.

#### Tumour tissue analyses

Details of tumour tissue analyses of MSI, BRAF, KRAS and CIMP have been reported previously.<sup>26</sup> In short, formalin-fixed, paraffin-embedded surgical specimens of CRC tumours were collected from cooperating pathology institutes and transferred to the tissue bank at the National Centre for Tumour Diseases (NCT) in Heidelberg. MSI analysis was performed using a mononucleotide marker panel (BAT25, BAT26 and CAT25), which differentiates MSI-high from non-MSI-high tumours with a sensitivity of 98.2% and a specificity of 100%, and with 100% concordance of MSI-high tumours compared with the National Cancer Institute/International Collaborative Group on HNPCC marker panel (BAT25, BAT26, D17S250, D2S123 and D5S346) for the evaluation of MSI in CRC.<sup>27–29</sup>

For KRAS, in about half of the tumour samples, mutation status was determined by a single-stranded conformational polymorphism technique using the same DNA sample, and expression of BRAF V600E was determined by immunohistochemical analyses by two pathologists independently (91% concordance,  $\kappa$  0.59). Discordant cases were discussed to obtain a final evaluation. In the other half of the tumour samples, KRAS mutation status and BRAF mutation status were determined by Sanger sequencing as reported previously.<sup>30</sup>

CIMP was determined after DNA bisulfite conversion as previously described.<sup>31</sup> CIMP-high and CIMP-low/negative were classified when 3–5 and 0–2 of the investigated loci (MGMT, MLH1, MINT1, MINT2, and MINT31) had a positive methylation status, respectively.

#### Statistical analyses

Multinomial logistic regression models were used to estimate adjusted odds ratios (ORs) and 95% confidence intervals (CIs) for the association of smoking and alcohol consumption with CRC risk according to molecular features and pathways. The models were adjusted for covariates known to be associated with CRC risk in all regression analyses: sex, age, body mass index (BMI) 5–14 years before diagnosis (cases) or interview (controls), education level, using non-steroidal anti-inflammatory drugs (NSAIDs) regularly for more than a year (yes, no), history of CRC in a first-degree family member, previous large bowel endoscopy and diabetes. Additionally, average lifetime daily ethanol consumption was included as a covariate in the smoking analyses and smoking (never, former, current) as a covariate in the alcohol consumption analyses. Ever, former and current smoking were compared to never smoking. High alcohol consumption was defined as the fourth quartile of the average daily lifetime gram ethanol consumption among alcohol drinkers (>24.6 g) and was compared in analyses to low/never consumption ( $\leq 24.6$  g).

In case-control analyses, each of the molecular features or pathways was compared to all study controls. In addition, combinations of single tumour markers approximating the traditional (MSS, CIMP-low/negative, BRAF-wt, KRAS-wt), sessile serrated (CIMP-high, BRAF-mut) and alternate (MSS, CIMP-low/negative, KRAS-mut) pathways to the development of CRC were examined.<sup>32</sup> To assess heterogeneity in CRC risk between subtypes and pathways, case-case analyses were conducted with the same covariates as in the case-control analyses. All statistical tests were two sided and the significance level ( $\alpha$ ) was <0.05. Analyses were conducted using R version 3.4.4.<sup>33</sup>

## RESULTS

A total of 4919 participants, 2444 cases and 2475 controls, were included in the current analysis. Descriptive statistics for study participants are shown in Table 1. Current smoking was associated with a 59% (OR = 1.59, 95% CI: 1.30–1.94) increased risk of CRC, while former smoking was associated with a 19% (OR = 1.19, 95% CI: 1.03–1.38) increased risk. However, risk was not increased if smoking cessation was more than 20 years ago. More than 29 pack years of smoking were associated with 61% increased CRC risk (OR = 1.61, 95% CI: 1.31–1.99) compared to never smoking. High alcohol consumption (>24.6 g/day) was associated with increased CRC risk (OR = 1.27, 95% CI: 1.08–1.50) (Supplementary Table 1).

#### Smoking and CRC risk by molecular pathological subtypes

Current smoking compared to never smoking showed much higher odd-ratios for MSI-high (OR = 2.79, 95% CI: 1.86–4.18) compared to MSS CRC (OR = 1.41, 95% CI: 1.14–1.75,  $p$ -heterogeneity = 0.001), for BRAF-mut (OR = 2.40, 95% CI: 1.41–4.07) compared to BRAF-wt CRC (OR = 1.52, 95% CI: 1.24–1.88,  $p$ -het = 0.074), for KRAS-wt (OR = 1.70, 95% CI: 1.36–2.13) compared to KRAS-mut CRC (OR = 1.26, 95% CI: 0.95–1.68,  $p$ -het = 0.039) and for CIMP-high CRC (OR = 2.01, 95% CI: 1.40–2.88) compared to CIMP-low/negative CRC (OR = 1.50, 95% CI: 1.22–1.85,  $p$ -het = 0.101), although not all differences in association were statistically significant at the  $p < 0.05$  level. The results and differences were not as strong for the comparison of ever vs. never smoking (Table 2 and Fig. 1).

#### Smoking and CRC risk by molecular pathological pathways

Ever and current smoking were significantly associated with higher risk of CRCs that were grouped into the traditional (OR = 1.50, 95% CI: 1.16–1.94) or the serrated pathways (OR = 2.39, 95%

**Table 1.** Characteristics of study population.

Variables	Cases (%), N = 2444	Controls (%), N = 2475	p Value <sup>a</sup>
<b>Gender</b>			
Female	1016 (41.6)	974 (40.4)	0.117
Male	1428 (58.4)	1501 (59.6)	
<b>Age (median (range))</b>			
	70 (30–96)	70 (34–99)	0.474
<b>BMI (kg/m<sup>2</sup>)</b>			
<25	758 (31.5)	928 (37.8)	<0.001
25–30	1168 (48.5)	1179 (48.1)	
>30	480 (20)	346 (14.1)	
<b>School education (years)</b>			
1–8	1672 (68.6)	1474 (59.7)	<0.001
9–10	403 (16.5)	485 (19.6)	
>10	363 (14.9)	511 (20.7)	
<b>Family history of CRC in first-degree relative</b>			
No	2068 (84.8)	2201 (89)	<0.001
Yes	370 (15.2)	271 (11)	
<b>Previous endoscopy</b>			
No	1904 (78.0)	1117 (45.1)	<0.001
Yes	538 (22.0)	1358 (54.9)	
<b>Diabetes</b>			
No	1989 (81.5)	2123 (85.9)	<0.001
Yes	452 (18.5)	348 (14.1)	
<b>Ever regular use of NSAIDs</b>			
Never	1866 (76.7)	1682 (68.4)	<0.001
Yes	568 (23.3)	778 (31.6)	
<b>Physical activity (metabolic equivalents MET-h/week)</b>			
Low	1122 (46.8)	1207 (47.1)	0.136
High	1277 (53.2)	1260 (52.9)	
<b>Smoking</b>			
Never	1134 (46.5)	1257 (50.9)	<0.001
Former	923 (37.9)	945 (38.3)	
Current	380 (15.6)	268 (10.9)	
<b>Avg. lifetime daily alcohol consumption (g ethanol)<sup>b</sup></b>			
None/low	1872 (76.9)	1996 (79.8)	0.012
High (>24.6 g)	562 (23.1)	473 (20.2)	

Cases and controls were matched by age and sex during recruitment to the study.  
BMI body mass index (kg/m<sup>2</sup>), NSAIDs non-steroidal anti-inflammatory drugs.  
<sup>a</sup>Fisher's exact test.  
<sup>b</sup>High alcohol consumption was defined as the fourth quartile of the average daily lifetime gram ethanol consumption among alcohol drinkers (>24.6 g) and was compared in analyses to low/never consumption (≤24.6 g).

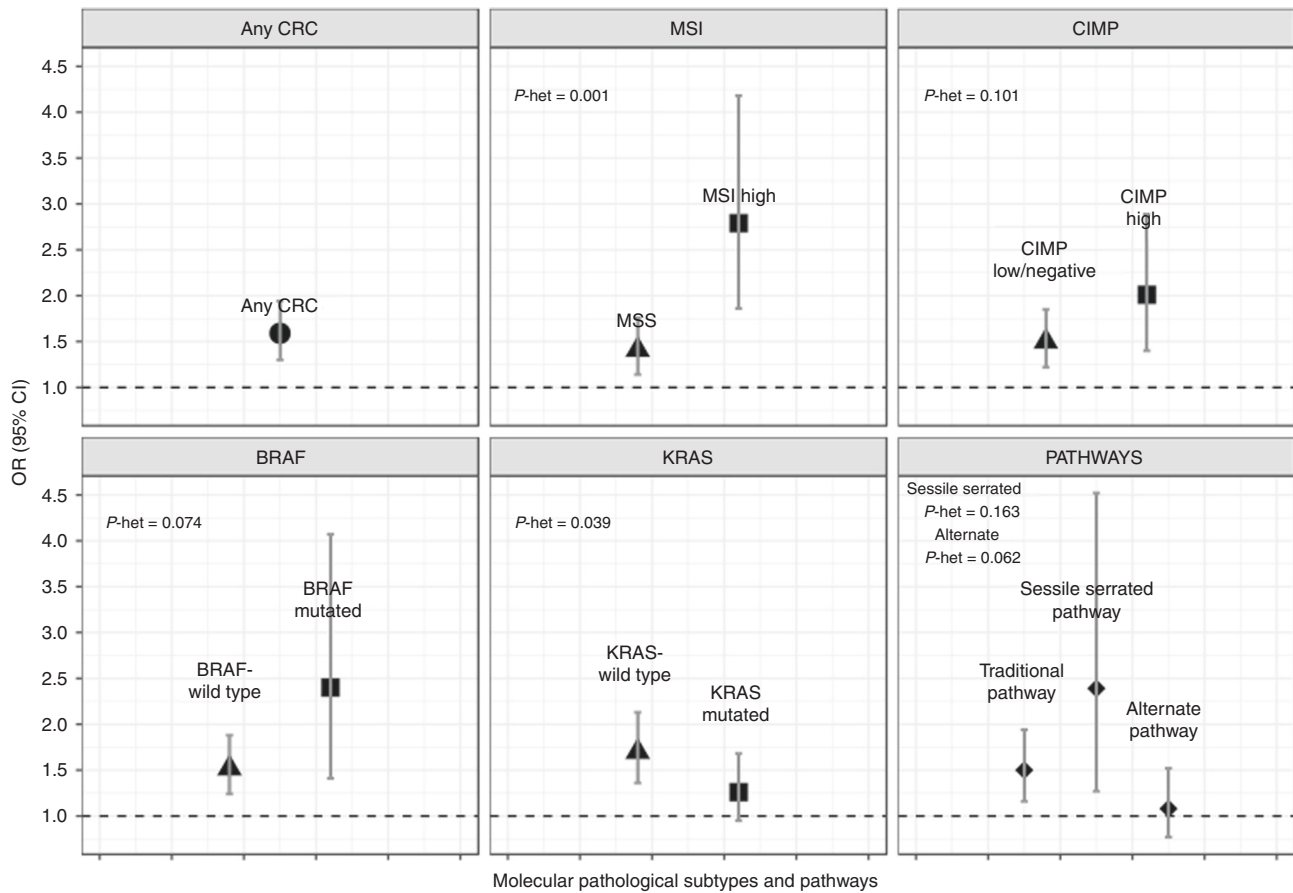
CI: 1.27–4.52) (Table 3 and Fig. 1). On the other hand, no association was found between smoking and CRC developing along the alternate pathway, characterised by MSS, CIMP-low/negative and KRAS mutation (OR = 1.08, 95% CI: 0.77–1.52, *p*-het = 0.062).

Alcohol consumption and CRC risk by molecular pathological subtypes and pathways  
Average lifetime daily consumption of more than 24.6 g ethanol was associated with around 30% increased CRC risk of the non-aberrant, more frequent subtypes (MSS, BRAF-wt, KRAS-wt, CIMP-low/neg) and not with the aberrant subtypes (MSI, BRAF-mut, CIMP-high), but no statistically significant differences were observed in heterogeneity testing between subtypes in this study. The strength of the associations of high alcohol consumption with

**Table 2.** Association of smoking and alcohol consumption with CRC risk by single molecular pathological subtypes.

	Ever regular smoking		Alcohol: g/ethanol/day <sup>a</sup>		OR (95% CI) <sup>b</sup>
	Never	Ever	≤24.6	>24.6	
	N (%)	N (%)	N (%)	N (%)	
Controls	1257 (50.8)	1217 (49.2)	1996 (80.8)	473 (19.2)	1
MSI	111 (47.0)	125 (53.0)	196 (82.7)	41 (17.3)	1.07 (0.72–1.58)
MSS	916 (46.8)	1041 (53.2)	1486 (76.1)	466 (23.9)	1.32 (1.11–1.57)
<i>p</i> -heterogeneity					0.385
BRAF-mut	91 (51.4)	86 (48.6)	152 (86.4)	24 (13.6)	1.09 (0.61–1.65)
BRAF-wt	955 (46.1)	1118 (53.9)	1574 (76)	496 (24)	1.30 (1.10–1.54)
<i>p</i> -heterogeneity					0.373
KRAS-mut	350 (47.9)	381 (52.1)	568 (77.9)	161 (22.1)	1.19 (0.94–1.49)
KRAS-wt	690 (45.6)	824 (54.4)	1147 (75.9)	364 (24.1)	1.30 (1.09–1.56)
<i>p</i> -heterogeneity					0.456
CIMP-high	196 (49.2)	202 (50.8)	323 (81)	76 (19)	1.22 (0.90–1.66)
CIMP-low/neg	930 (46.1)	1088 (53.9)	1529 (76)	484 (24)	1.28 (1.08–1.52)
<i>p</i> -heterogeneity					0.868
Former					
			OR (95% CI) <sup>b</sup>		
			N (%)		
Current					
			OR (95% CI) <sup>b</sup>		
			N (%)		
Controls	945 (38.3)	268 (10.9)	1	1	
MSI	79 (33.6)	45 (19.1)	2.79 (1.86–4.18)	2.79 (1.86–4.18)	
MSS	756 (38.6)	285 (14.6)	1.41 (1.14–1.75)	1.41 (1.14–1.75)	
<i>p</i> -heterogeneity			0.584	0.001	
BRAF-mut	63 (35.6)	23 (13.0)	1.56 (1.07–2.28)	2.40 (1.41–4.07)	
BRAF-wt	793 (38.3)	324 (15.6)	1.17 (1.01–1.36)	1.52 (1.24–1.88)	
<i>p</i> -heterogeneity			0.075	0.074	
KRAS-mut	285 (39.0)	96 (13.1)	1.17 (0.96–1.43)	1.26 (0.95–1.68)	
KRAS-wt	577 (38.1)	246 (16.3)	1.22 (1.04–1.44)	1.70 (1.36–2.13)	
<i>p</i> -heterogeneity			0.226	0.039	
CIMP-high	147 (36.9)	55 (13.8)	1.32 (1.02–1.71)	2.01 (1.40–2.88)	
CIMP-low/neg	771 (38.2)	316 (15.7)	1.18 (1.01–1.37)	1.50 (1.22–1.85)	
<i>p</i> -heterogeneity			0.128	0.101	

<sup>a</sup>Multinomial logistic regression model adjusted for: Sex, age, BMI, education level, history of colorectal cancer in first-degree relative, previous endoscopy, diabetes, ever NSAIDs regular use and average lifetime alcohol consumption/ever regular smoking. Ever/former/current smoking compared to never smoking.  
<sup>b</sup>High alcohol consumption was defined as the fourth quartile of the average daily lifetime gram ethanol consumption among alcohol drinkers (>24.6 g) and was compared in analyses to low/never consumption (≤24.6 g).



**Fig. 1 Association of current smoking with CRC risk overall and by molecular pathological subtypes and pathways.** Heterogeneity between molecular subtypes was assessed in case-case comparison. For the pathways, heterogeneity was assessed using the traditional pathway as reference.

**Table 3.** Association of smoking and alcohol consumption with CRC risk by molecular pathological pathways.

	Regular smoking		OR (95% CI) <sup>b</sup>	Alcohol: g/ethanol/day <sup>a</sup>							
	Never	Ever		Former		Current		≤24.6		>24.6	
	N (%)	N (%)		N (%)	OR (95% CI) <sup>b</sup>	N (%)	OR (95% CI) <sup>b</sup>	N (%)	N (%)	OR (95% CI) <sup>b</sup>	
Controls	1235 (51.1)	1183 (49.4)	1	916 (37.9)	1	267 (11.0)	1	1953 (80.8)	465 (19.2)	1	
Traditional pathway	414 (44.5)	517 (55.5)	1.30 (1.09–1.54)	368 (39.5)	1.23 (1.02–1.49)	149 (16.0)	1.50 (1.16–1.94)	680 (73.0)	251 (27.0)	1.57 (1.16–2.11)	
Sessile serrated pathway	65 (53.7)	56 (46.3)	1.55 (1.01–2.36)	41 (33.9)	1.34 (0.85–2.13)	15 (12.4)	2.39 (1.27–4.52)	101 (83.5)	20 (16.5)	1.84 (0.75–4.52)	
<i>p</i> -het traditional vs. sessile serrated			0.340		0.604		0.163			0.850	
Alternate pathway	254 (49.4)	260 (50.6)	1.11 (0.9–1.39)	197 (38.3)	1.13 (0.89–1.43)	63 (12.3)	1.08 (0.77–1.52)	397 (77.2)	117 (22.8)	1.15 (0.80–1.66)	
<i>p</i> -het traditional vs. alternate			0.171		0.425		0.062			0.243	

<sup>a</sup>Logistic regression model adjusted for: Sex, age, BMI, education level, history of colorectal cancer in first-degree relative, previous endoscopy, diabetes, ever NSAIDs regular use and average lifetime alcohol consumption/ever regular smoking. Ever/former/current smoking compared to never smoking. High alcohol intake compared to low/none. Traditional pathway: MSS, CIMP-low/negative, BRAF-wt and KRAS-wt; sessile serrated pathway: CIMP-high, BRAF-mut; alternate pathway: MSS, CIMP-low/negative, KRAS-mut.

<sup>b</sup>High alcohol consumption was defined as the third quartile of the average daily lifetime gram ethanol consumption among alcohol drinkers (>24.6 g) and was compared in analyses to low/never consumption (≤24.6 g).

risk of traditional pathway CRC and the direction of association of the serrated pathway CRC were comparable to that of current smoking, but heterogeneity was not statistically significant (Tables 2 and 3 and Fig. 2).

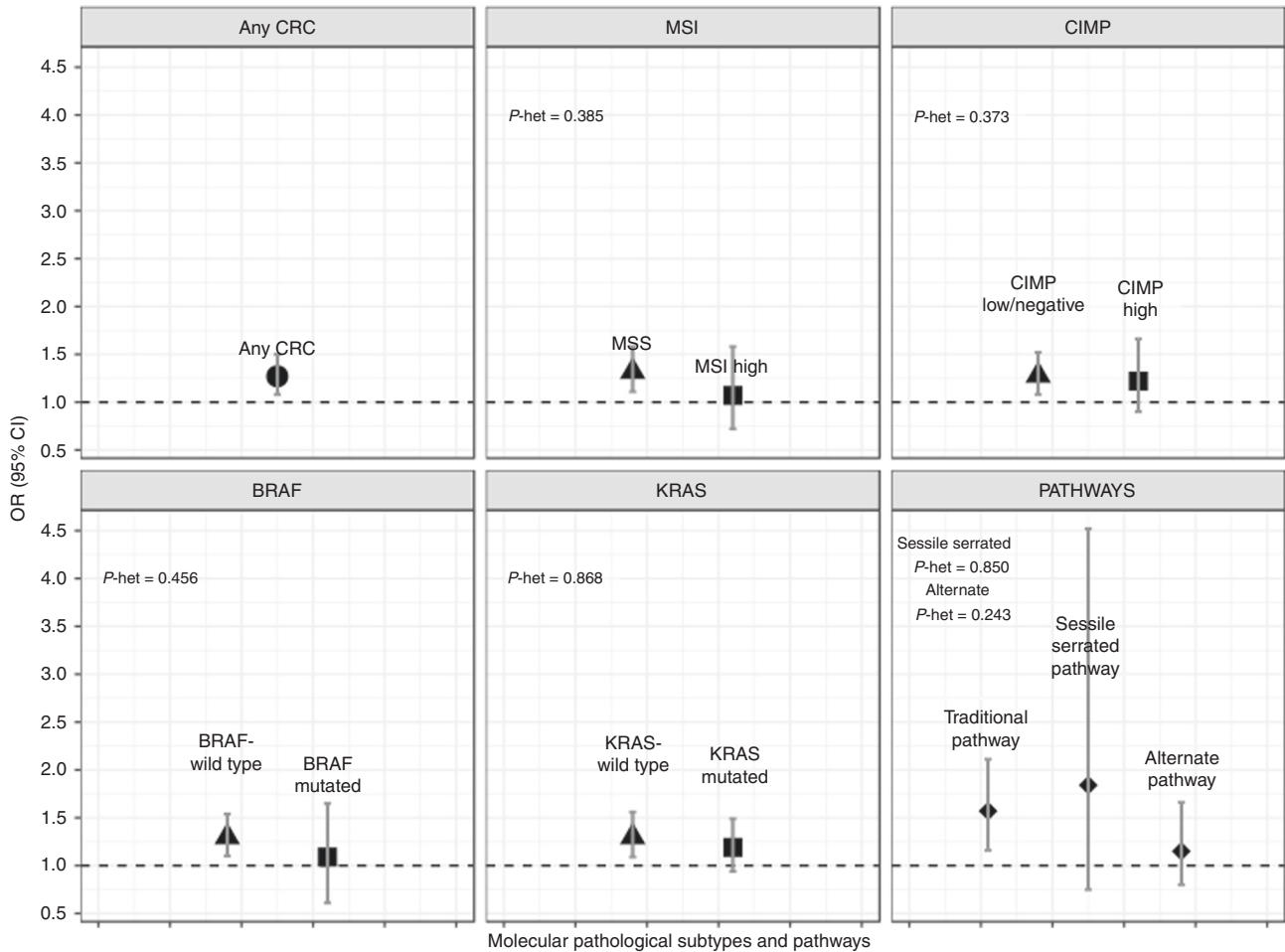
**Joint effects of smoking and alcohol**

No interaction was found between high lifetime average daily alcohol consumption and ever smoking (*p*-interaction = 0.519) in the association with CRC risk, and no major differences were

found when analysing alcohol consumption (high vs. low/none) stratified by smoking status (ever/never) and CRC risk for the different molecular subtypes or pathways (data not shown).

**DISCUSSION**

This large population-based case-control study aimed to examine the association between smoking and alcohol consumption and CRC risk by CRC subtypes and pathways characterised by MSI,



**Fig. 2 Association of high alcohol consumption with CRC risk overall and by molecular pathological subtypes and pathways.** Heterogeneity between molecular subtypes was assessed in case-case comparison. For the pathways, heterogeneity was assessed using the traditional pathway as reference.

BRAF mutation, KRAS mutation and CIMP status. Ever and current smoking were associated with higher risk for CRC, especially with MSI-high, BRAF-mut, KRAS-wt and CIMP-high CRC. Smoking was also associated with higher risk of cancers developing via the traditional or the serrated pathways. High alcohol consumption was not differently associated with single CRC subtypes or pathways, but associations with molecular pathways still seemed similar to those of current smoking.

Our findings are generally in agreement with former studies. Smoking was associated with higher MSI-high compared to MSS CRC risk in five previous studies<sup>6–10</sup> and in a meta-analysis published in 2018.<sup>34</sup> Stronger associations were also previously reported for BRAF-mut compared to BRAF-wt CRC,<sup>8,10,11</sup> for KRAS-wt compared to KRAS-mut CRC<sup>12–14</sup> and for CIMP-high compared to CIMP-low/negative CRC.<sup>10,11</sup> Smoking has also been found to be associated with the serrated-polyps pathway, defined by CIMP-high and BRAF-mut status.<sup>35–37</sup> Further, in accordance with previous studies, no major or statistically significant differences were found in the associations between alcohol consumption and CRC risk by molecular pathological subtypes,<sup>9,17–19,22</sup> although the observed associations with CRC pathways pointed to potential differences in our study.

Although several possible biological mechanisms were proposed,<sup>10,38</sup> the way smoking increases the risk of MSI-high CRC is still not established. Our results also support a strong link between smoking and BRAF mutation, which is regarded as the initiating

event in sessile serrated adenomas, followed by methylation of key tumour suppressor genes, which would also be supported by the stronger association with CIMP-high CRC and the sessile serrated pathway.<sup>32,39</sup>

Molecular pathological epidemiology (MPE) focuses on heterogeneous aetiology of CRC based on molecular tumour features. While studies linking smoking with CRC risk found an overall increase of around 26% in risk,<sup>40,41</sup> our study provides more specific risk estimates by known CRC subtypes and pathways. This distinction can help provide more evidence for the causal relationship and its mechanisms between smoking and CRC risk.

The large size of the study, its population-based design, the comprehensive assessment of smoking, alcohol consumption and other lifestyle, medical and family history factors, and the analysis of multiple major molecular tumour tissue markers, are notable strengths of this study. To the best of our knowledge, this is the first study examining the effect of both smoking and alcohol consumption on CRC risk by major molecular pathological tumour features and pathways of CRC. This study adds to the limited knowledge about the potential smoking-related increase in CRC risk according to molecular features, potentially linking smoking differentially to specific molecular pathways.

The study also has limitations. MPE is a relatively new field of research, thus analyses are often exploratory and need confirmation from other studies.<sup>1</sup> Not all patients with available tumour tissue samples could be included in the subtype analyses: MSI status was

available for 90% of cases, BRAF for 92%, KRAS for 92% and CIMP for 99% of cases. Another limitation is that some analysed subgroups of cases were small, which lowered the ability of the study to reach significant results, particularly when multiple tumour features were combined. As this is an observational study, based on self-reports during standardised interviews, smoking, alcohol and other relevant factors may be subject to information bias.

In summary, based on results from this large population-based case-control study, smoking, and in particular current smoking, showed the strongest association with increased risk of molecular subtypes of CRC MSI-high and KRAS-wt and with CRC showing features of the sessile serrated pathway. No major differences were observed for the association of alcohol with subtypes, but potential differences according to pathways should be investigated in future studies. More large studies with tumour marker combinations are needed to confirm these results for a better characterisation of the carcinogenic mechanisms underlying these associations.

## ACKNOWLEDGEMENTS

We thank Ute Handte-Daub, Ansgar Brandhorst, Petra Bächer, Bettina Walter, Gloria Laukemper and Terence Osere for their excellent technical assistance. We thank the study participants and the interviewers who collected the data. We also thank the following hospitals and cooperating institutions that recruited patients for this study: Chirurgische Universitätsklinik Heidelberg, Klinik am Gesundbrunnen Heilbronn, St. Vincentiuskrankenhaus Speyer, St. Josefskrankenhaus Heidelberg, Chirurgische Universitätsklinik Mannheim, Diakonissenkrankenhaus Speyer, Krankenhaus Salem Heidelberg, Kreiskrankenhaus Schwetzingen, St. Marienkrankenhaus Ludwigshafen, Klinikum Ludwigshafen, Stadtklinik Frankenthal, Diakoniekrankenhaus Mannheim, Kreiskrankenhaus Sinsheim, Klinikum am Plattenwald Bad Friedrichshall, Kreiskrankenhaus Weinheim, Kreiskrankenhaus Eberbach, Kreiskrankenhaus Buchen, Kreiskrankenhaus Mosbach, Enddarmzentrum Mannheim, Kreiskrankenhaus Brackenheim and Cancer Registry of Rhineland-Palatinate, Mainz. We are also very grateful for the support of the pathologies in the provision of tumour samples: Institut für Pathologie, Universitätsklinik Heidelberg; Institut für Pathologie, Klinikum Heilbronn; Institut für Angewandte Pathologie, Speyer; Pathologisches Institut, Universitätsklinikum Mannheim; Institut für Pathologie, Klinikum Ludwigshafen; Institut für Pathologie, Klinikum Stuttgart; Institut für Pathologie, Klinikum Ludwigsburg. Special thanks to the tissue bank of National Centre for Tumour Diseases (NCT), Heidelberg, for storage and processing of the tissue samples.

## AUTHOR CONTRIBUTIONS

Study concepts and design: E.L.A., M.H., H.Br., J.C.-C.; data acquisition: L.J., E.A., W.R., E. H., M.K., H.Bl., J.C.-C., M.H., H.Br.; data analysis and interpretation: E.L.A., P.R.C., M.H., L. J.; manuscript preparation: E.L.A., M.H.; manuscript review and approval: E.L.A., P.R.C., L.J., E.A., W.R., M.K., H.Bl., J.C.-C., H.Br., M.H.; funding acquisition: J.C.-C., M.H., H.Br.

## ADDITIONAL INFORMATION

**Ethics approval and consent to participate** The DACHS study was approved by the ethics committees of the Medical Faculty of Heidelberg University and the state medical boards of Baden-Wuerttemberg and Rhineland-Palatinate. Written informed consent for study participation and publication of results based on aggregated data was obtained from all participants. The study was performed in accordance with the Declaration of Helsinki.

**Data availability** The data used in the current analyses are not publicly available due to legal privacy restrictions.

**Competing interests** The authors declare no competing interests.

**Funding information** This work was supported by the German Research Council (BR 1704/6-1, BR 1704/6-3, BR 1704/6-4, BR 1704/6-6, CH 117/1-1, HO 5117/2-1, HE 5998/2-1, KL 2354/3-1, RO 2270/8-1 and BR 1704/17-1); The German Federal Ministry of Education and Research (01KH0404, 01ER0814, 01ER0815, 01ER1505A and 01ER1505B); The Interdisciplinary Research Programme of the National Centre for Tumour Diseases (NCT), Germany. Open access funding provided by Projekt DEAL.

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41416-020-0803-0>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

- Ogino, S., Chan, A. T., Fuchs, C. S. & Giovannucci, E. Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. *Gut* **60**, 397–411 (2011).
- Carethers, J. M. & Jung, B. H. Genetics and genetic biomarkers in sporadic colorectal cancer. *Gastroenterology* **149**, 1177–1190 (2015). e3.
- Morkel, M., Riemer, P., Blaker, H. & Sers, C. Similar but different: distinct roles for KRAS and BRAF oncogenes in colorectal cancer development and therapy resistance. *Oncotarget* **6**, 20785–20800 (2015).
- Tsoi, K. K., Pau, C. Y., Wu, W. K., Chan, F. K., Griffiths, S. & Sung, J. J. Cigarette smoking and the risk of colorectal cancer: a meta-analysis of prospective cohort studies. *Clin. Gastroenterol. Hepatol.* **7**, 682–688 (2009). e1–5.
- Fedirko, V., Tramacere, I., Bagnardi, V., Rota, M., Scotti, L., Islami, F. et al. Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies. *Ann. Oncol.* **22**, 1958–1972 (2011).
- Slattery, M. L., Curtin, K., Anderson, K., Ma, K. N., Ballard, L., Edwards, S. et al. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. *J. Natl Cancer Inst.* **92**, 1831–1836 (2000).
- Chia, V. M., Newcomb, P. A., Bigler, J., Morimoto, L. M., Thibodeau, S. N. & Potter, J. D. Risk of microsatellite-unstable colorectal cancer is associated jointly with smoking and nonsteroidal anti-inflammatory drug use. *Cancer Res.* **66**, 6877–6883 (2006).
- Curtin, K., Samowitz, W. S., Wolff, R. K., Herrick, J., Caan, B. J. & Slattery, M. L. Somatic alterations, metabolizing genes and smoking in rectal cancer. *Int. J. Cancer* **125**, 158–164 (2009).
- Poynter, J. N., Haile, R. W., Siegmund, K. D., Campbell, P. T., Figueiredo, J. C., Limburg, P. et al. Associations between smoking, alcohol consumption, and colorectal cancer, overall and by tumor microsatellite instability status. *Cancer Epidemiol. Biomark. Prev.* **18**, 2745–2750 (2009).
- Limsui, D., Vierkant, R. A., Tillmans, L. S., Wang, A. H., Weisenberger, D. J., Laird, P. W. et al. Cigarette smoking and colorectal cancer risk by molecularly defined subtypes. *J. Natl Cancer Inst.* **102**, 1012–1022 (2010).
- Samowitz, W. S., Albertsen, H., Sweeney, C., Herrick, J., Caan, B. J., Anderson, K. E. et al. Association of smoking, CpG island methylator phenotype, and V600E BRAF mutations in colon cancer. *J. Natl Cancer Inst.* **98**, 1731–1738 (2006).
- Samadder, N. J., Vierkant, R. A., Tillmans, L. S., Wang, A. H., Lynch, C. F., Anderson, K. E. et al. Cigarette smoking and colorectal cancer risk by KRAS mutation status among older women. *Am. J. Gastroenterol.* **107**, 782–789 (2012).
- Slattery, M. L., Anderson, K., Curtin, K., Ma, K., Schaffer, D., Edwards, S. et al. Lifestyle factors and Ki-ras mutations in colon cancer tumors. *Mutat. Res.* **483**, 73–81 (2001).
- Li, W., Qiu, T., Ling, Y., Guo, L., Li, L. & Ying, J. Molecular pathological epidemiology of colorectal cancer in Chinese patients with KRAS and BRAF mutations. *Oncotarget* **6**, 39607–39613 (2015).
- Wark, P. A., Van der Kuil, W., Ploemacher, J., Van Muijen, G. N., Mulder, C. J., Weijenberg, M. P. et al. Diet, lifestyle and risk of K-ras mutation-positive and -negative colorectal adenomas. *Int. J. Cancer* **119**, 398–405 (2006).
- Diergaarde, B., Vrieling, A., van Kraats, A. A., van Muijen, G. N., Kok, F. J. & Kampman, E. Cigarette smoking and genetic alterations in sporadic colon carcinomas. *Carcinogenesis* **24**, 565–571 (2003).
- Slattery, M. L., Anderson, K., Curtin, K., Ma, K. N., Schaffer, D. & Samowitz, W. Dietary intake and microsatellite instability in colon tumors. *Int. J. Cancer* **93**, 601–607 (2001).
- Diergaarde, B., Braam, H., van Muijen, G. N., Ligtenberg, M. J., Kok, F. J. & Kampman, E. Dietary factors and microsatellite instability in sporadic colon carcinomas. *Cancer Epidemiol. Biomark. Prev.* **12**(Part 1), 1130–1136 (2003).
- Schernhammer, E. S., Giovannucci, E., Baba, Y., Fuchs, C. S. & Ogino, S. B vitamins, methionine and alcohol intake and risk of colon cancer in relation to BRAF mutation and CpG island methylator phenotype (CIMP). *PLoS ONE* **6**, e21102 (2011).
- Slattery, M. L., Curtin, K., Sweeney, C., Levin, T. R., Potter, J., Wolff, R. K. et al. Diet and lifestyle factor associations with CpG island methylator phenotype and BRAF mutations in colon cancer. *Int. J. Cancer* **120**, 656–663 (2007).
- de Vogel, S., Bongaerts, B. W., Wouters, K. A., Kester, A. D., Schouten, L. J., de Goeij, A. F. et al. Associations of dietary methyl donor intake with MLH1 promoter hypermethylation and related molecular phenotypes in sporadic colorectal cancer. *Carcinogenesis* **29**, 1765–1773 (2008).

22. Bongaerts, B. W., de Goeij, A. F., van den Brandt, P. A. & Weijnenberg, M. P. Alcohol and the risk of colon and rectal cancer with mutations in the K-ras gene. *Alcohol* **38**, 147–154 (2006).
23. Brenner, H., Chang-Claude, J., Seiler, C. M., Rickert, A. & Hoffmeister, M. Protection from colorectal cancer after colonoscopy: a population-based, case-control study. *Ann. Intern. Med.* **154**, 22–30 (2011).
24. Hoffmeister, M., Jansen, L., Rudolph, A., Toth, C., Kloor, M., Roth, W. et al. Statin use and survival after colorectal cancer: the importance of comprehensive confounder adjustment. *J. Natl Cancer Inst.* **107**, djv045 (2015).
25. Scherz, H. & Souci, S. W. (eds). *Food Composition and Nutrition Tables* 5th edn (Medpharm Scientific Publishers, Stuttgart, 1994).
26. Amitay, E. L., Carr, P. R., Jansen, L., Walter, V., Roth, W., Herpel, E. et al. Association of aspirin and nonsteroidal anti-inflammatory drugs with colorectal cancer risk by molecular subtypes. *J. Natl Cancer Inst.* **22**, 475–483 (2019). <https://doi.org/10.1093/jnci/djy170>
27. Hoffmeister, M., Blaker, H., Kloor, M., Roth, W., Toth, C., Herpel, E. et al. Body mass index and microsatellite instability in colorectal cancer: a population-based study. *Cancer Epidemiol. Biomark. Prev.* **22**, 2303–2311 (2013).
28. Findeisen, P., Kloor, M., Merx, S., Sutter, C., Woerner, S. M., Dostmann, N. et al. T25 repeat in the 3' untranslated region of the CASP2 gene: a sensitive and specific marker for microsatellite instability in colorectal cancer. *Cancer Res.* **65**, 8072–8078 (2005).
29. Boland, C. R., Thibodeau, S. N., Hamilton, S. R., Sidransky, D., Eshleman, J. R., Burt, R. W. et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res.* **58**, 5248–5257 (1998).
30. Jia, M., Jansen, L., Walter, V., Tagscherer, K., Roth, W., Herpel, E. et al. No association of CpG island methylator phenotype and colorectal cancer survival: population-based study. *Br. J. Cancer* **115**, 1359–1366 (2016).
31. Warth, A., Kloor, M., Schirmacher, P. & Blaker, H. Genetics and epigenetics of small bowel adenocarcinoma: the interactions of CIN, MSI, and CIMP. *Mod. Pathol.* **24**, 564–570 (2011).
32. Leggett, B. & Whitehall, V. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology* **138**, 2088–2100 (2010).
33. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2015).
34. Carr, P. R., Alwers, E., Bienert, S., Weberpals, J., Kloor, M., Brenner, H. et al. Lifestyle factors and risk of sporadic colorectal cancer by microsatellite instability status: a systematic review and meta-analyses. *Ann. Oncol.* **29**, 825–834 (2018).
35. Phipps, A. I., Limburg, P. J., Baron, J. A., Burnett-Hartman, A. N., Weisenberger, D. J., Laird, P. W. et al. Association between molecular subtypes of colorectal cancer and patient survival. *Gastroenterology* **148**, 77–87 (2015). e2.
36. Baillie, L., Loughrey, M. B. & Coleman, H. G. Lifestyle risk factors for serrated colorectal polyps: a systematic review and meta-analysis. *Gastroenterology* **152**, 92–104 (2017).
37. He, X., Wu, K., Ogino, S., Giovannucci, E. L., Chan, A. T. & Song, M. Association between risk factors for colorectal cancer and risk of serrated polyps and conventional adenomas. *Gastroenterology* **155**, 355–373.e18 (2018). <https://doi.org/10.1053/j.gastro.2018.04.019>
38. Yang, P., Cunningham, J. M., Halling, K. C., Lesnick, T. G., Burgart, L. J., Wiegert, E. M. et al. Higher risk of mismatch repair-deficient colorectal cancer in alpha(1)-antitrypsin deficiency carriers and cigarette smokers. *Mol. Genet. Metab.* **71**, 639–645 (2000).
39. Bettington, M., Walker, N., Clouston, A., Brown, I., Leggett, B. & Whitehall, V. The serrated pathway to colorectal carcinoma: current concepts and challenges. *Histopathology* **62**, 367–386 (2013).
40. Gong, J., Hutter, C., Baron, J. A., Berndt, S., Caan, B., Campbell, P. T. et al. A pooled analysis of smoking and colorectal cancer: timing of exposure and interactions with environmental factors. *Cancer Epidemiol. Biomark. Prev.* **21**, 1974–1985 (2012).
41. Johnson, C. M., Wei, C., Ensor, J. E., Smolenski, D. J., Amos, C. I., Levin, B. et al. Meta-analyses of colorectal cancer risk factors. *Cancer Causes Control* **24**, 1207–1222 (2013).



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.