BJC British Journal of Cancer

www.nature.com/bjc

ARTICLE

Clinical Study

Association between risk factors, molecular features and CpG island methylator phenotype colorectal cancer among different age groups in a Taiwanese cohort

Kuo-Hsing Chen[1,2,3], Liang-In Lin[4,5], Chang-Tsu Yuan[6], Li-Hui Tseng[7], Yu-Liang Chao[2], Yi-Hsin Liang[1,2,3], Jin-Tung Liang[8], Been-Ren Lin[8], Ann-Lii Cheng[1,2,3] and Kun-Huei Yeh[2,3]

BACKGROUND: CpG island methylator phenotype (CIMP) represents a carcinogenesis pathway of colorectal cancer (CRC) and the association between CIMP CRC, molecular features and risk factors in East Asian population is less studied.

METHODS: We prospectively enrolled newly diagnosed CRC patients at the National Taiwan University Hospital. Clinicopathological data and risk factors for CRC were collected during interview. The tumour samples were subjected to CIMP, *RAS/BRAF* mutation and microsatellite instability tests. CIMP-high was determined when ≧3 methylated loci of *p16, MINT1, MINT2, MINT31* and *MLH1* were identified. Multivariate logistic regression was used to evaluate the association between risk factors and CIMP-high CRC.

RESULTS: Compared with CIMP-low/negative CRC, CIMP-high CRC was associated with more stage IV disease, *BRAF V600E* mutation and high body mass index (BMI ≧ 27.5 kg/m^2) in younger patients (age < 50 y), and more right-sided tumour, *BRAF V600E* mutation, MSI-high and colorectal polyp in elder patients (age ≧ 50 y). Multivariate analyses showed that BMI ≧27.5 kg/m^2 was significantly associated with CIMP-high CRC in younger patients.

CONCLUSIONS: We identified distinct clinicopathological features for CIMP-high CRC among different age groups in Taiwan. Our data suggest the association between BMI ≧27.5 kg/m^2 and CIMP-high CRC in patients younger than 50 years.

British Journal of Cancer (2021) 125:48–54; https://doi.org/10.1038/s41416-021-01300-5

BACKGROUND

Colorectal cancer (CRC) is the third most common cancer worldwide, with ~1800,000 newly diagnosed cases and estimated 860,000 cancer-related deaths per year in 2018.[1] In the past three decades, the incidence of CRC in Taiwan has been rapidly increasing (Fig. 1).[2] In 2017, the age-adjusted incidence rate of CRC in male and female was 52.83 and 35.11 per 100,000 people. The age-specific rate was getting higher along with increase of age, specifically, 43.29 per 100,000 people in age 45–49 years, 156 per 100,000 people in age 60–64 years and 411 per 100,0000 in age 80–84 years. Thus, defining the risk factors for CRC is an important research issue.

CpG island methylator phenotype (CIMP) represents a unique pattern of epigenomic instability in CRC, characterised by wild-spread DNA hypermethylation in multiple promoter CpG islands leading to the silence of hundreds of genes per cancer cell.[3,4] CIMP-high CRC accounts for ~15% of all CRC cases and has been reported to correlate with old age, female sex, *BRAF V600E* mutation, sporadic mismatch repair deficient CRC, mucinous histology and proximal colon cancer.[5–8] In one large cohort study conducted by Weisenberger et al., there was a difference in association between CRC risk factors and CRC with CIMP-high or CIMP-low/negative.[5] In detail, there was no association between family history of CRC and CIMP-high CRC, a positive association between body mass index (BMI) (kg/m^2), pack-years of smoking, and CIMP-high CRC in female patients, and a positive association between non-steroidal anti-inflammatory drugs (NSAIDs) use in all patients, a negative association between alcohol consumption and CIMP-high CRC in female patients and a negative association between men and women who engaged in higher levels of physical activities, compared to CIMP-low/negative CRC. One meta-analysis also revealed that there was no association between CIMP-high CRC and family history.[8] However, whether this phenomenon is also valid in the East Asian Taiwanese population remains unknown.

Weisenberger's study also revealed that the percentage of CIMP-high CRC in young CRC patients (age < 50 years) was only around 5%, which was significantly lower than that in older CRC patients (age ≥ 70 years, 19.5%).[5] In our previous study, we showed that in Taiwan, CIMP-high CRC represented 16.4% of CRC cases.[9] Notably, we found that the frequency of CIMP-high CRC in young CRC patients (age < 50 years) was in the same range as that in elderly patients (14.3% vs. 17%). In this study, we aimed to

[1]Department of Medical Oncology, National Taiwan University Cancer Center, Taipei, Taiwan; [2]Department of Oncology, National Taiwan University Hospital, Taipei, Taiwan; [3]Graduate Institute of Oncology, National Taiwan University College of Medicine, Taipei, Taiwan; [4]Department of Laboratory Medicine, National Taiwan University Hospital, Taipei, Taiwan; [5]Department of Clinical Laboratory Sciences and Medical Biotechnology, National Taiwan University College of Medicine, Taipei, Taiwan; [6]Department of Pathology, National Taiwan University Hospital, Taipei, Taiwan; [7]Department of Medical Genetics, National Taiwan University Hospital, Taipei, Taiwan and [8]Department of Surgery, National Taiwan University Hospital, Taipei, Taiwan
Correspondence: Kun-Huei Yeh (khyeh@ntu.edu.tw)

Received: 6 September 2020 Revised: 27 January 2021 Accepted: 3 February 2021
Published online: 12 April 2021

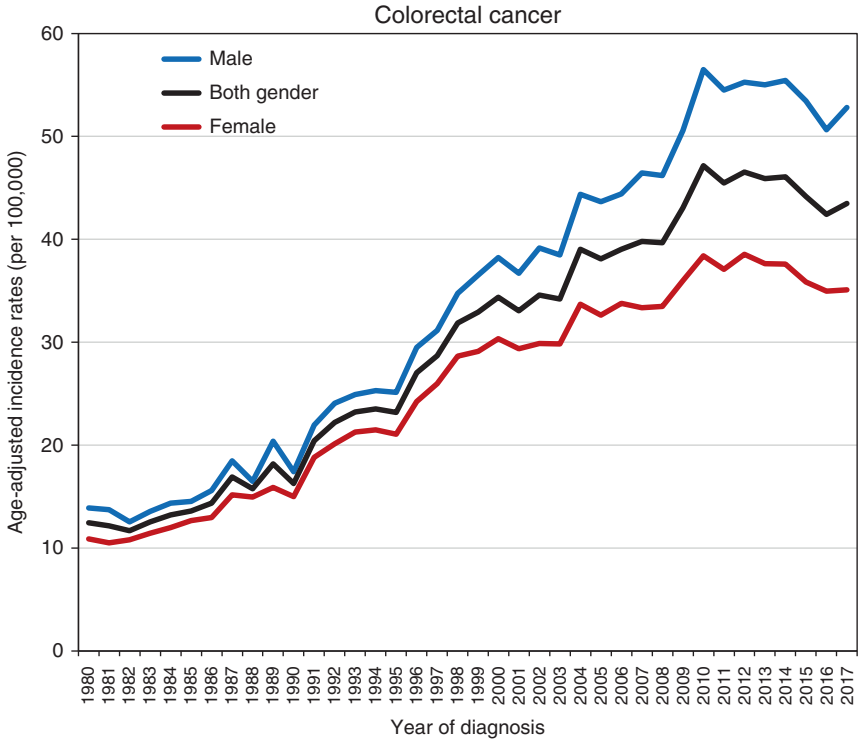

Fig. 1 Age-adjusted incidence rates (ASR) of colorectal cancer (ICD10 codes C18-C20) for both sexes in Taiwan using the WHO standard population 2000. Data have been obtained from Taiwan Cancer Registry Statistical Service (1980–2017). ICD10 International Classification of Disease 10th revision, WHO World Health Organization.

validate the high frequency of CIMP-high CRC in young patients in Taiwan and explore the association between CRC risk factors, molecular features and CIMP phenotype in different age groups.

METHODS

Patient eligibility

We initially planned to enrol 320 CRC patients in our prospective cohort. In order to focus on young CRC patients, we aimed to include at least 96 patients (30%) who were younger than 50 years at the time of diagnosis. Patients were eligible for this study if they met the following inclusion criteria: (1) they had a pathology-or cytology-confirmed CRC, and had computed tomography for staging CRC (American Joint Cancer Committee on Cancer (AJCC) system, 7th edition) (2) they were 20 years old or older and (3) they planned to undergo a colectomy. Patients who had received chemotherapy and radiotherapy for CRC before colectomy were excluded. Informed consent and interviews with patients were performed before obtaining the tissue specimens and written informed consent was obtained from all patients. We collected the patients' formalin-fixed paraffin-embedded (FFPE) tissues as well as clinicopathological characteristics, such as age, sex, stage at diagnosis, primary tumour site, histology, tumour grade and BMI and the risk factors for CRC. The primary tumour site was designated as the left-sided colon if located from the descending colon to the sigmoid colon, the right-sided colon if located from the cecum to the splenic flexure. Histological data, including the presence or absence of a mucinous component, medullary histology, or signet-ring cell histology and grade, were obtained by microscopic examination. Mucinous carcinoma was determined if there were presence of mucinous component more than 50% of total tumour volume. Tumours were graded as high or low grade. The tumour grade was low if more than 95% of tumour harbouring glandular formation or high-degree of microsatellite instability. Otherwise, it was high-grade. Risk factor items for CRC included family history, BMI at diagnosis, colorectal polyp, diabetes mellitus (DM), hyperlipidaemia, non-steroidal anti-inflammatory drug (NSAID) use and hormone replacement therapy (HRT) before CRC diagnosis. Collected tumour samples were subjected to *KRAS/NRAS* exon 2, 3 and 4 mutations and *BRAF V600E* mutation analyses, microsatellite instability (MSI) and CIMP analyses. This study was approved by The Institutional Review Board of the National Taiwan University Hospital (NTUH).

Risk factors and clinical data

Risk factor data were collected by interviewing the patients and checking the electronic medical records at NTUH. Patients were categorised into two groups according to age: <50, and ≥50 years. Patients were considered to have a family history of CRC if they reported that at least one or more first-degree family members had CRC. BMI at diagnosis was determined based on the medical records. BMI was categorised to three groups according to WHO criteria as 18–24.9, 25–29.9 and ≧30. BMI was also categorised to two groups based on the opinion reported by WHO expert consultation for appropriate BMI for Asian populations: <27.5 and ≧27.5. In that report, they identified BMI 27.5 or higher as representing high risk for public health action.[10] History of colorectal polyps, DM, hyperlipidaemia, pre-diagnosis NSAIDs and HRT use were obtained from patient interviews and electronic medical records. The criteria used for diagnosing DM (type 1 or type 2) and hyperlipidaemia was based on self-report and/or reported use of medications for these conditions. NSAIDs use was coded as "yes" if the patient had ever used NSAIDs ≧2 times per week for 1 month or longer within 1 year. HRT use was coded as "yes" if the patient had ever used any hormone replacement preparation for 6 months or longer within one year.

Methylation analysis

Genomic DNA was extracted using a QIAamp DNA FFPE Tissue Kit (Qiagen, Santa Clarita, CA, USA). Following the recommendation of

the EZ DNA Methylation Kit protocol (Zymo Research, Irvine, CA, USA), genomic DNA was treated with sodium bisulphite. Finally, we used a 5-gene panel (*p16, MINT1, MINT2, MINT31,* and *MLH1*) and the MethyLight assay to determine CIMP status. The detailed method of CIMP analysis and the primers and probes used for the MethyLight study have been described in our previous study.[9] The percentage of methylated reference was calculated by an equation described previously and loci with percentage of methylated reference values >10 were regarded as methylated.[11] CIMP-high CRC was designated when three or more (≥3) methylated loci were identified, and CIMP-low/negative CRC was identified when fewer than three (<3) methylated loci were found.

RAS, BRAF mutational and microsatellite instability analysis

The genomic DNA of the samples was subjected to real-time polymerase chain reaction (PCR). Primer pairs that encompassed exons 2, 3 and 4 (codon 12, 13 and 61) of *KRAS*, *NRAS* and exon 15 of *BRAF* were used. Direct sequencing was performed using an automated ABI 3730 sequencer (Applied Biosystems, Foster City, CA, USA). Detailed methods on *KRAS, NRAS* and *BRAF* mutational analyses have been described.[9,12] Microsatellite instability status was determined by multiplex PCR amplification of five mononucleotide microsatellite loci (*BAT-25, BAT-26, NR-21, NR-24* and *MONO-27*), which was performed using the MSI Analysis System (Promega Corp., Madison, WI, USA). The neoplasm was designated as MSI-high if novel allele lengths were identified in the neoplastic cells compared with the normal/germline at two or more (≥2) of five microsatellite loci.[13]

Statistical analysis

Patients' baseline characteristics were summarised as frequencies and percentages. The association between CIMP and clinical characteristics was estimated by using logistic regression (Table 1), Chi-square test, and Fisher's exact test (Table 2 and Supplement Table S1). The association between CIMP and risk factors were estimated by logistic regression (Table 3 and Supplement Table S2). Multivariate logistic regression model was used to evaluate the odds ratios of variables for CIMP-high CRC. A 2-sided *p*-value less than 0.05 was considered statistically significant. Only the variables whose *p*-values less than 0.1 in Table 2 and Table 3 were included in multivariate logistic regression model. SAS, version 9.4, statistical software (SAS Institute, Cary, NC, USA) was used for all statistical analyses.

RESULTS

A total of 99 patients younger than 50 years and 221 patients older than 50 years were enrolled in the prospective cohort, between March 2016 and June 2019. All clinicopathological variables and risk factors for CRC were recorded. Among the tumour samples, 293 (91.6%) were adequate for CIMP analysis, 288 (90%) for *RAS* mutation, 286 (89.4%) for *BRAF V600E* mutation and 318 (99.4%) for MSI tests. The most common reason for the lack of analysis in some tumour samples was an insufficient amount of DNA for analysis. The sample size in the primary analysis was 293, including 89 (30.4%) early onset CRC.

There was no missing methylation on any marker during CIMP calculation. The frequency of CIMP-high CRC was 15.4% in the prospective cohort. The distribution of the clinicopathological variables in CIMP-high and CIMP-low/negative CRC is shown in Table 1. Briefly, CIMP-high CRC was associated with right-sided colon as the primary site ($p = 0.03$), *BRAF V600E* mutation ($p < 0.01$) and MSI-high tumours ($p < 0.01$), but NOT with older age ($p = 0.91$) and female ($p = 0.49$). The distribution of the clinicopathological variables in CIMP-high and CIMP-low/negative CRC according to different age groups (<50 years and ≥50 years) is shown in Table 2. In this analysis, *BRAF V600E* mutation was significantly associated with CIMP-high CRC in both groups, while stage IV disease at diagnosis was significantly associated with CIMP-high CRC in early onset (age < 50 years) CRC and right-sided colon as the primary site and MSI-high tumours were significantly more common in late onset (age ≥ 50 years) CRC.

The frequencies of CIMP-high CRC among the two age groups were 15.7% (14/89) in patients aged <50 years, and 15.2% (31/204) in patients aged ≥50 years. The result showing a high frequency of early onset CIMP-high CRC was consistent with data from our previous retrospective cohort ($N = 450$), which showed that the

Table 1. Distribution of CIMP status by clinicopathologic variables.

Variable, *N* (%)	CIMP-high CRC ($N = 45$)	CIMP-low/ negative CRC ($N = 248$)	OR (95% CI) high vs. low/negative	*p*-value
Age (years)				
<50	14 (31.1)	75 (30.2)	Reference	
≥50	31 (68.9)	173 (69.8)	0.96 (0.48–1.91)	0.91
Sex				
Male	22 (48.9)	135 (54.4)	Reference	
Female	23 (51.1)	113 (45.6)	1.25 (0.66–2.36)	0.49
Stage[a]				
1	7 (15.6)	41 (16.5)	Reference	
2	8 (17.8)	72 (29.0)	0.65 (0.22–1.93)	0.44
3	17 (37.8)	96 (37.1)	1.08 (0.42–2.81)	0.87
4	13 (28.9)	39 (17.3)	1.77 (0.64–4.88)	0.27
Histology				
Mucinous adenocarcinoma	3 (6.8)	8 (3.2)	Reference	
Not mucinous adenocarcinoma	42 (93.3)	240 (96.8)	0.41 (0.10–1.64)	0.21
Tumour grade				
Low	39 (86.7)	230 (92.7)	Reference	
High	5 (11.1)	13 (5.2)	2.27 (0.77–6.72)	0.14
Missing	1 (2.2)	5 (2.0)	NA	
Primary tumour side[b]				
Left-sided	18 (40.0)	126 (50.8)	Reference	
Rectum	5 (11.1)	51 (20.6)	0.69 (0.24–1.95)	0.48
Right-sided	22 (48.9)	71 (28.6)	2.17 (1.09–4.31)	**0.03**
RAS				
Wild-type	28 (62.2)	158 (63.7)	Reference	
KRAS Mutation	16 (35.5)	83 (33.5)	1.09 (0.56–2.12)	0.81
NRAS Mutation	1 (2.2)	2 (0.8)	2.82 (0.25–32.17)	0.40
Missing	0 (0.0)	5 (2.0)	NA	
BRAF V600E				
Wild-type	26 (57.8)	238 (96.0)	Reference	
Mutation	18 (40.0)	4 (1.6)	41.19 (12.96–130.95)	**<0.01**
Missing	1 (2.2)	6 (2.4)	NA	
MSI				
Low/stable	31 (68.9)	229 (93.1)	Reference	
High	14 (31.1)	17 (6.9)	6.08 (2.73–13.55)	**<0.01**
Missing	0 (0.0)	2 (0.8)	NA	

Bold values indicate statistical significance $p < 0.05$.
Logistic regression was used for statistical analysis.
CIMP CpG island methylator phenotype, *OR* odds ratio, *CI* confidence interval, *MSI* microsatellite instability.
[a]By the American Joint Cancer Committee on Cancer (AJCC) system, 7th edition.
[b]The right-sided colon is defined as the cecum to the splenic flexure of the colon. The left-sided colon is defined as the region from the descending colon to sigmoid colon.

Table 2. Distribution of CIMP status by clinicopathologic variables stratified by age.

Variable, *N* (%)	Age < 50			Age ≧ 50		
	CIMP-high CRC (*N* = 14)	CIMP-low/negative CRC (*N* = 75)	*p*-value	CIMP-high CRC (*N* = 31)	CIMP-low/negative CRC (*N* = 173)	*p*-value
Sex						
Male	8 (57.1)	47 (62.7)		14 (45.2)	88 (50.9)	
Female	6 (42.9)	28 (37.3)	0.70	17 (54.8)	85 (49.1)	0.56
Stage[b]						
1–3	5 (35.7)	59 (74.7)		27 (87.1)	149 (86.1)	
4	9 (64.3)	16 (25.3)	**<0.01**[a]	4 (12.9)	24 (13.9)	1.00[a]
Histology						
Mucinous adenocarcinoma	1 (7.1)	2 (2.7)		2 (6.5)	5 (2.9)	
Not Mucinous adenocarcinoma	13 (92.9)	73 (97.3)	0.41[a]	29 (93.6)	167 (97.1)	0.29[a]
Tumour grade						
Low	11 (78.6)	69 (92.0)		28 (90.3)	161 (93.1)	
High	2 (14.3)	4 (5.3)	0.22[a]	3 (9.7)	9 (5.2)	0.40[a]
Missing	1 (7.1)	2 (2.7)		0 (0.0)	3 (1.7)	
Primary tumour side[c]						
Left-sided + rectum	9 (64.3)	60 (80.0)		14 (45.2)	117 (67.6)	
Right-sided	5 (35.7)	15 (20.0)	0.29[a]	17 (54.8)	56 (32.4)	**0.02**
RAS						
Wild-type	9 (64.3)	50 (66.7)		19 (61.3)	108 (62.4)	
Mutation	5 (35.7)	23 (30.7)	0.76[a]	12 (38.7)	62 (35.8)	0.81
Missing	0 (0.0)	2 (2.7)		0 (0.0)	3 (1.7)	
BRAF V600E						
Wild-type	8 (57.1)	71 (94.7)		18 (58.1)	167 (96.5)	
Mutation	6 (42.9)	2 (2.7)	**<0.01**[a]	12 (38.7)	2 (1.2)	**<0.01**[a]
Missing	0 (0.0)	2 (2.7)		0 (0.0)	3 (1.7)	
MSI						
Low/stable	11 (78.6)	69 (92.0)		20 (64.5)	160 (93.6)	
High	3 (21.4)	6 (8.0)	0.15[a]	11 (35.5)	11 (6.4)	**<0.01**[a]
Missing	0 (0.0)	0 (0.0)		0 (0.0)	2 (1.2)	

Bold values indicate statistical significance $p < 0.05$.
Chi-square test, and Fisher's exact test were used for statistical analysis.
CIMP CpG island methylator phenotype, *OR* odds ratio, *CI* confidence interval, *MSI* microsatellite instability.
[a]Fisher's exact test.
[b]By the American Joint Cancer Committee on Cancer (AJCC) system, 7th edition.
[c]The right-sided colon is defined as the cecum to the splenic flexure of the colon. The left-sided colon is defined as the region from the descending colon to sigmoid colon.

frequency of CIMP-high CRC in patients who were age <50 years was 14.3% (11 of 77).[9]

The association of CRC risk factors with CIMP-high and CIMP-low/negative phenotypes in overall patients and patients with early onset and late onset CRC is shown in Table 3. In this analysis, we found that patients with early onset CIMP-high CRC had a trend for more BMI 27.5 kg/m^2 or higher ($p = 0.09$). On the other hand, patients with late onset CIMP-high CRC had significant more colorectal polyp ($p = 0.03$). The sensitivity analysis of CIMP-high CRC, molecular features and risk factors by gender was also performed and the results were shown in Supplement Table S1 and S2. The results showed *BRAF V600E* mutation and MSI-high were associated with CIMP-high CRC in both male and female patients, and the association between CIMP-high CRC and *RAS* mutation (in male), CIMP-high CRC and right-sided tumour and colorectal polyp (in female) were noted in different subgroups.

In multivariate logistic regression model, BMI 27.5 kg/m^2 or higher ($p = 0.03$) and stage IV disease at diagnosis ($p < 0.01$) were independent variables for early onset CIMP-high CRC (Table 4), while in patients with late onset CRC, MSI-high ($p < 0.01$) was the only independent variable for CIMP-high CRC (Table 4).

DISCUSSION

In this study, we showed a positive association between CIMP-high CRC and stage IV disease at diagnosis, right-sided colon cancer, *BRAF V600E* mutation and MSI-high tumours, which is consistent with the findings of our previous retrospective cohort and the literature.[6,11,14] The findings that were different from the analyses in Western populations were the negative association between CIMP-high CRC and old age and female. Through the stratification by age, we also found the association between clinicopathological characteristics and CIMP-high CRC was different in early onset and late onset CRC.

Table 3. Distribution of CIMP status by risk factors of CRC in overall patients and stratified by age.

Variable, *N* (%)	Overall				Age < 50				Age ≧ 50			
	CIMP-high CRC (*N* = 45)	CIMP-low/negative CRC (*N* = 248)	OR (95% CI) High vs. low/negative	*p*-value	CIMP-high CRC (*N* = 14)	CIMP-low/negative CRC (*N* = 75)	OR (95% CI) High vs. low/negative	*p*-value	CIMP-high CRC (*N* = 31)	CIMP-low/negative CRC (*N* = 173)	OR (95% CI) High vs. low/negative	*p*-value
Family history												
None	42 (93.3)	218 (87.9)	Reference		13 (92.9)	68 (90.7)	Reference		29 (93.55)	150 (86.7)	Reference	
Yes	3 (6.7)	30 (12.1)	0.52 (0.15–1.78)	0.30	1 (7.1)	7 (9.3)	0.75 (0.09–6.60)	0.79	2 (6.45)	23 (13.3)	0.45 (0.10–2.01)	0.30
Colorectal polyp												
0	32 (71.1)	209 (85.3)	Reference		12 (85.7)	70 (93.3)	Reference		20 (64.5)	139 (81.8)	Reference	
≥1	13 (28.9)	36 (14.7)	2.36 (1.13–4.92)	**0.02**	2 (14.3)	5 (6.7)	2.33 (0.41–13.43)	0.34	11 (35.5)	31 (18.2)	2.47 (1.07–5.67)	**0.03**
BMI at diagnosis (kg/m^2), WHO criteria												
<25	29 (64.4)	162 (65.7)	Reference		7 (50.0)	51 (68.0)	Reference		22 (71.0)	112 (64.7)	Reference	
25–29	14 (31.1)	70 (28.2)	1.12 (0.56–2.26)	0.74	6 (42.9)	18 (24.0)	2.43 (0.72–8.19)	0.15	8 (25.8)	52 (30.1)	0.78 (0.33–1.88)	0.58
≥30	2 (4.4)	15 (6.1)	0.75 (0.16–3.45)	0.71	1 (7.1)	6 (8.0)	1.21 (0.13–11.63)	0.87	1 (3.2)	9 (5.2)	0.57 (0.07–4.69)	0.60
BMI at diagnosis (kg/m^2)												
<27.5	37 (82.2)	205 (82.7)	Reference		9 (64.3)	63 (84.0)	Reference		28 (90.3)	142 (82.1)	Reference	
≥27.5	8 (17.8)	43 (17.3)	1.03 (0.45–2.37)	0.94	5 (35.7)	12 (16.0)	2.92 (0.83–10.24)	0.09	3 (9.7)	31 (17.9	0.49 (0.14–1.72)	0.27
DM												
None	37 (82.2)	201 (81.4)	Reference		14 (100.0)	72 (96.0)	Reference		23 (74.2)	129 (75.0)	Reference	
Yes	8 (17.8)	46 (18.6)	0.95 (0.41–2.16)	0.89	0 (0.0)	3 (4.0)	NA	0.98	8 (25.8)	43 (25.0)	1.04 (0.43–2.51)	0.92
Hyperlipidemia												
None	31 (68.9)	189 (76.2)	Reference		13 (92.9)	64 (85.3)	Reference		18 (58.1)	125 (72.25)	Reference	
Yes	14 (31.1)	59 (23.8)	1.45 (0.72–2.90)	0.30	1 (7.1)	11 (14.7)	0.45 (0.05–3.77)	0.46	13 (41.9)	48 (27.75)	1.88 (0.86–4.13)	0.12
NSAIDs use												
None	41 (91.1)	234 (94.7)	Reference		13 (92.9)	75 (100.0)	Reference		28 (90.3)	159 (92.4)	Reference	
Yes	4 (8.9)	13 (5.3)	1.76 (0.55–5.65)	0.35	1 (7.1)	0 (0.0)	NA	0.99	3 (9.7)	13 (7.6)	1.31 (0.35–4.90)	0.69
HRT												
None	45 (100.0)	243 (98.4)	Reference		14 (100.0)	75 (100.0)	Reference		31 (100.0)	168 (97.7)	Reference	
Yes	0 (0.0)	4 (1.6)	NA	NA	0 (0.0)	0 (0.0)	NA	NA	0 (0.0)	4 (2.3)	NA	0.98

Bold values indicate statistical significance $p < 0.05$.
Logistic regression was used for statistical analysis.
CIMP CpG island methylator phenotype, *OR* odds ratio, *CI* confidence interval, *BMI* body mass index, *DM* diabetes mellitus, *NSAIDs* non-steroidal anti-inflammatory drugs, *HRT* hormone replacement therapy, *WHO* world health organization.

Table 4. Multivariate logistic regression for odds ration of CIMP-high CRC in overall patients, patients with age <50 y and age ≥50 y.

Variable, *N* (%)	Overall		Age < 50 y		Age ≥ 50 y	
	OR (95% CI) High vs. low/negative	*p*-value	OR (95% CI) High vs. low/negative	*p*-value	OR (95% CI) High vs. low/negative	*p*-value
Primary tumour site as right-sided	1.55 (0.67–3.60)	0.30			1.62 (0.69–3.82)	0.27
MSI-high	5.64 (2.06–15.45)	**<0.01**			5.86 (2.07–16.60)	**<0.01**
Colorectal polyp	1.33 (0.49–3.61)	0.58			1.38 (0.54–3.58)	0.50
BRAF V600E	38.86 (11.60–130.12)	**<0.01**				
Stage IV at diagnosis			10.021 (2.430–41.326)	**<0.01**		
BMI (≥27.5 kg/m^2)			5.657 (1.211–26.418)	**0.03**		

Bold values indicate statistical significance $p < 0.05$.
BRAF V600E mutation was excluded in multivariate logistic regression in patients with age <50 y and age ≥50 y because of few events were noted.
CIMP CpG island methylator phenotype, *OR* odds ratio, *CI* confidence interval, *BMI* body mass index.

In addition, we demonstrated that there was a high frequency (around 15%) of CIMP-high CRC in young (age < 50 y) CRC patients. These findings are in line with report (17.9%, based on the definition of ≧3/5 Weisenberger's panel) in young CRC patients (age < 50 y) from a Chinese cohort,[15] and were different from those of the analysis of a large cohort in Western countries, where the frequency of CIMP-high CRC in young (age < 50 y) patients was around 5%.[5] Although the CIMP panel used in Weisenberger's study was different from that used in our study, a study has demonstrated that there was no significant difference in the correlation of clinicopathological factors in the two panels.[16] These results reveal that there may be different frequencies for CIMP-high CRC in young patients between East Asian and Western populations. Further studies are warranted to confirm this phenomenon in other East Asian countries.

In our study, we also showed that BMI 27.5 kg/m^2 or higher were independently associated with CIMP-high CRC in patients who were younger than 50 years in multivariate analysis. However, because of small sample size in early onset CIMP-high CRC, so this finding needs further validation. In the Weisenberger's study, increase BMI (kg/m^2) was associated with CIMP-high CRC in female patients,[5] but this phenomenon was not found in our study (Supplementary table S2). There had been increasing evidence to show increased cardiovascular diseases and high prevalence of type 2 diabetes in Asian countries where the average BMI was below 25 kg/m^2, which was the cut-off point that defined overweight in WHO criteria.[10] The WHO expert consultation then provided recommendations in 2002, as adding the trigger points of 23, 27.5, 32.5 and 37.5 kg/m^2 for points for public health action and a BMI 27.5 kg/m^2 or higher represented the high risk.[10] Thus, our study highlights the importance of maintaining BMI less than 27.5 kg/m^2, especially in those aged <50 years.

The reason behind the correlation between CIMP-high CRC and a higher BMI in patient who were younger than 50 years is intriguing. The correlation of CIMP-high CRC and a higher BMI may be related to the biologic effects of excessive calorie intake, instead of restriction. In a Dutch Hunger Winter cohort study, early life exposure to famine (calorie restriction) correlated with decreased risk of CRC, specifically, CIMP-high CRC.[17] Another study revealed that caloric restriction delayed the methylation drift during aging in mammals.[18] In addition, accelerated epigenetic aging has been proposed to be associated with multiple types of cancer, including CRC.[19,20] Epigenetic aging has been developed from genome-wide DNA methylation array platforms and refers to the cumulative work done by an epigenetic maintenance system.[19] CIMP-high has been shown to be significantly associated with epigenetic aging in CRC and obesity may accelerate epigenetic aging in human tissues.[19,21,22] Thus, there may be more accumulated age-related methylations in overweight and obese patients and it contributes to heavily methylated tumours. But why there was only significant association of BMI in young patients (<50 years) with CIMP-high CRC but not in elder patients (≥50 years)? According to a recent report, increased BMI was associated with accelerated epigenetic aging in the middle-aged individuals (40–49 years), but not in individuals aged in 15–24 years and >90 years.[23] Instead, in very old individuals, they observed the higher BMI corresponded to lower epigenetic age. Thus, we think that higher BMI accelerates epigenetic aging overtime and it needs long enough time to have obesity-induced epigenetic aging in colon and rectum, which leads to cancer in middle-age. But in elder CRC patients, they have exposed environmental factors for decades that would cause various genetic and epigenomic changes, so the effects of obesity-induced epigenetic aging may not be influential.

Limitations of this study include: (1) the relatively small sample size, thus the association of variables and CIMP in different age and gender groups might be attributed to sample size; (2) the lack of the reliable information about smoking and alcohol consumption and (3) there were few NSAIDs and HRT users that made the analysis inconclusive. There have been multiple studies demonstrated the association between smoking and CIMP-high CRC and there seemed to be dose-response relationship in smoking and the risk of CIMP-high CRC, thus the lack of the information about smoking may be influential[5,7,24,25] On the other hand, previous studies showed conflicting results regarding the association between alcohol consumption and CIMP-high CRC and this issue remains inconclusive.[5,26] Strengths of this study include the fact that it is a prospective study and the fact that we were able to collect data, such as BMI and the other risk factors. This study is also the first study to evaluate the association between CRC risk factors and CIMP in one of the East Asian population. We also report that although associated clinicopathological features are the same in the Taiwanese cohort and the Western populations, the association of CRC risk factors (age, sex and BMI) with CIMP-high CRC may vary among different populations.

CONCLUSIONS

This study demonstrates the distinct clinicopathological features and risk factors of CIMP-high CRC among different age groups in Taiwan. Our data validate the finding that there is a high frequency of early onset CIMP-high CRC in Taiwan and suggest high BMI (27.5 or higher) may be associated to these tumours.

ACKNOWLEDGEMENTS

We are grateful for the statistical assistance and calculation of age-adjusted incidence rates (ASR) of colorectal cancer provided by the Department of Medical Research at National Taiwan University Hospital.

AUTHOR CONTRIBUTIONS

K.H.C. and K.H.Y. designed the study. L.I.L. performed CIMP and *RAS/BRAF* analyses. L.H.T. performed MSI analysis. K.H.C., Y.L.C., Y.H.L., J.T.L. and B.R.L. collected patient's clinical information and surgical specimens. C.T.Y. handled surgical specimens and did pathology interpretation. K.H.C. did data analysis. K.H.C., A.L.C. and K.H.Y. interpreted data and wrote manuscript.

ADDITIONAL INFORMATION

Ethics approval and consent to participate This study was approved by The Institutional Review Board of the National Taiwan University Hospital (NTUH).

Consent for publication Not applicable.

Data availability The datasets generated and/or analysed during the current study are not publicly available because another study is ongoing but are available from the corresponding author on reasonable request.

Competing interests The authors declare no competing interests.

Funding information This work was supported by the Ministry of Health and Welfare, R.O.C. (Taiwan) (DOH100-TD-C-111-001).

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41416-021-01300-5.

REFERENCES

1. Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A. & Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**, 394–424 (2018).
2. Health Promotion Administration, Ministry of Health Welfare. *Taiwan Cancer Registry Statistical Service (1980–2017)*. http://tcr.cph.ntu.edu.tw/main.php?Page=A5B2(Chinese). (2020).
3. Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J. G., Baylin, S. B. & Issa, J. P. CpG island methylator phenotype in colorectal cancer. *Proc. Natl Acad. Sci. USA* **96**, 8681–8686 (1999).
4. Issa, J. P. CpG island methylator phenotype in cancer. *Nat. Rev. Cancer* **4**, 988–993 (2004).
5. Weisenberger, D. J., Levine, A. J., Long, T. I., Buchanan, D. D., Walters, R., Clendenning, M. et al. Association of the colorectal CpG island methylator phenotype with molecular features, risk factors, and family history. Cancer epidemiology, biomarkers & prevention: a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive. *Oncology* **24**, 512–519 (2015).
6. Barault, L., Charon-Barra, C., Jooste, V., de la Vega, M. F., Martin, L., Roignot, P. et al. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. *Cancer Res.* **68**, 8541–8546 (2008).
7. 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).
8. Advani, S. M., Advani, P., DeSantis, S. M., Brown, D., VonVille, H. M., Lam, M. et al. Clinical, pathological, and molecular characteristics of CpG island methylator phenotype in colorectal cancer: a systematic review and meta-analysis. *Transl. Oncol.* **11**, 1188–1201 (2018).
9. Chen, K. H., Lin, L. I., Tseng, L. H., Lin, Y. L., Liau, J. Y., Tsai, J. H. et al. CpG island methylator phenotype may predict poor overall survival of patients with stage IV colorectal cancer. *Oncology* **96**, 156–163 (2019).
10. WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet* **363**, 157–163 (2004).
11. Weisenberger, D. J., Siegmund, K. D., Campan, M., Young, J., Long, T. I., Faasse, M. A. et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat. Genet.* **38**, 787–793 (2006).
12. De Roock, W., Claes, B., Bernasconi, D., De Schutter, J., Biesmans, B., Fountzilas, G. et al. Effects of KRAS, BRAF, NRAS, and PIK3CA mutations on the efficacy of cetuximab plus chemotherapy in chemotherapy-refractory metastatic colorectal cancer: a retrospective consortium analysis. *Lancet Oncol.* **11**, 753–762 (2010).
13. 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).
14. Samowitz, W. S., Albertsen, H., Herrick, J., Levin, T. R., Sweeney, C., Murtaugh, M. A. et al. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology* **129**, 837–845 (2005).
15. Liu, J., Tang, L., Yi, J., Li, G., Lu, Y., Xu, Y. et al. Unique characteristics of CpG island methylator phenotype (CIMP) in a Chinese population with colorectal cancer. *BMC Gastroenterol.* **19**, 173 (2019).
16. Lee, S., Cho, N. Y., Yoo, E. J., Kim, J. H. & Kang, G. H. CpG island methylator phenotype in colorectal cancers: comparison of the new and classic CpG island methylator phenotype marker panels. *Arch. Pathol. Lab. Med.* **132**, 1657–1665 (2008).
17. Hughes, L. A., Simons, C. C., van den Brandt, P. A., Goldbohm, R. A., de Goeij, A. F., de Bruine, A. P. et al. Body size, physical activity and risk of colorectal cancer with or without the CpG island methylator phenotype (CIMP). *PLoS ONE* **6**, e18571 (2011).
18. Maegawa, S., Lu, Y., Tahara, T., Lee, J. T., Madzo, J., Liang, S. et al. Caloric restriction delays age-related methylation drift. *Nat. Commun.* **8**, 539 (2017).
19. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
20. Wang, T., Maden, S. K., Luebeck, G. E., Li, C. I., Newcomb, P. A., Ulrich, C. M. et al. Dysfunctional epigenetic aging of the normal colon and colorectal cancer risk. *Clin. Epigenet.* **12**, 5 (2020).
21. Horvath, S., Erhart, W., Brosch, M., Ammerpohl, O., von Schonfels, W., Ahrens, M. et al. Obesity accelerates epigenetic aging of human liver. *Proc. Natl Acad. Sci. USA* **111**, 15538–15543 (2014).
22. de Toro-Martín, J., Guénard, F., Tchernof, A., Hould, F.-S., Lebel, S., Julien, F. et al. Body mass index is associated with epigenetic age acceleration in the visceral adipose tissue of subjects with severe obesity. *Clin. Epigenet.* **11**, 172 (2019).
23. Nevalainen, T., Kananen, L., Marttila, S., Jylhava, J., Mononen, N., Kahonen, M. et al. Obesity accelerates epigenetic aging in middle-aged but not in elderly individuals. *Clin. Epigenet.* **9**, 20 (2017).
24. 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).
25. Nishihara, R., Morikawa, T., Kuchiba, A., Lochhead, P., Yamauchi, M., Liao, X. et al. A prospective study of duration of smoking cessation and colorectal cancer risk by epigenetics-related tumor classification. *Am. J. Epidemiol.* **178**, 84–100 (2013).
26. Amitay, E. L., Carr, P. R., Jansen, L., Roth, W., Alwers, E., Herpel, E. et al. Smoking, alcohol consumption and colorectal cancer risk by molecular pathological subtypes and pathways. *Br. J. Cancer* **122**, 1604–1610 (2020).