Supplementary Table S1. Details of populations for model development and validation

Author, year	Country	Period of study	Selection of cases	Selection of controls	Exclusions	Number of cases:controls	Gender of cases:controls (male, %)	Age of cases:controls (mean (SD))
Abe 2017	Japan	2001-2005 (d) 2005-2013 (v)	Histologically confirmed CRC at one hospital	Randomly selected from patients attending outpatient clinic during the same period with no history of cancer or neoplasms confirmed by examination. Matched by age (+/- 5 years) and sex.	Prior history of cancer	558:1116 (d) 547:547 (v)	62.7:62.7 (d) 65.5:65.5 (v)	60 (10.3): 60 (10.10) (d) 61 (9.96): 61 (9.96) (v)
Dunlop 2013a and b	UK, Canada, Australia, USA and Germany (d) Sweden and Finland (v)	Prior to 2013 (d and v)	Hospitals, clinical genetics centres and oncology clinics records from 13 datasets for development and two datasets for validation	Healthy individuals and cancer-free spouses (d) and blood donors and cancer-free spouses (v)		24,395: 19,994 (d) 2479:2169 (v)	56.7:46.6 (d) Not reported (v)	N/A
Frampton 2016	UK	2001-2012 (v)	CRC cases from VQ58 GWAS (two UK-based adjuvant chemotherapy clinical trials (VICTOR and QUASAR2))	From UK population-based 1958 Birth Cohort	Not given	1800:2690	Not reported	Not reported
Hosono 2016a and b	Japan	2001-2005 (d) 2005-2013 (v)	Histologically confirmed CRC at one hospital	Randomly selected from patients attending outpatient clinic during the same period with no history of cancer or neoplasms confirmed by examination. Matched by age (+/- 5 years) and sex.	Prior history of cancer	558:1116 (d) 547:547 (v)	62.7:62.7 (d) 65.5:65.5 (v)	60 (10.3): 60 (10.10) (d) 61 (9.96): 61 (9.96) (v)
Hsu 2015	USA and Germany	1990-2011 (d) 1993-2001 (v)	Colorectal adenocarcinoma confirmed by medical records, pathology reports, or death certificate from 2 case-control and 4 nested case-control studies and 1 screening trial for validation	Randomly selected from population registries, drivers' license records or state identification lists.	Not of European descent	5811:6302 (d) 866:869 (v)	Not reported	Not reported
Huyghe 2019	European (91.7%) and East Asian (8.3%) (d)	Not given (d)	Varied between incorporated studies.	Varied between incorporated studies.	Not given	58131:67347	Not reported	Not reported
Ibanez-Sanz 2017a and b	Spain	2008-2013 (d and v)	Age 20-85 with histologically confirmed CRC, having resided in the catchment area for at least 6 months prior to recruitment. Enrolled in 23 hospitals and primary care centres in 12 Spanish provinces.	Frequency-matched to cases, by age, sex and region, ensuring that in each region there was at least one control of the same sex and 5-year interval for each case.	Personal history of CRC	1336:2744	53.53:64.75	Not reported
Iwasaki 2017	Japan	1990-2009 (d and v)	New cases of CRC in the JPHC cohort study (9 public health centres) identified through active patient notifications, data linkage with	Individuals in JPHC cohort study with no prior history of CRC at time when case diagnosed. Matched by sex, age (within 3 years), date of	Women, history of cancer at baseline or no blood sample	341:329	100:100	Not reported

			population-based cancer registries and death certificates	blood withdrawal (within 3 months), time since last meal at blood withdrawal (within 4 hours), and study location				
Jenkins 2016, 2019	Australia, Canada, USA	N/A (v)	Simulation study: Simulation population including 1,000,000 individuals with CRC External validation study: invasive colon or rectum cancer identified from population-based cancer registries.	Simulation study: Simulation population including 1,000,000 individuals without CRC External validation study: persons with no CRC diagnosis randomly selected from the general population by using Medicare and Driver's License files, telephone subscriber lists, or electoral rolls.	N/A External validation study: germline mutations in the DNA mismatch repair genes and MUTYH	N/A External validation study: 1181:999	N/A External validation study: 52.0:47.9	N/A External validation study: 53.0 (11.4):59.9 (11.0)
Jeon 2018 a and b	Australia, Canada, Germany, Israel and USA.	1992-2005 (d and v)	Varied between incorporated studies.	Varied between incorporated studies.	CRC cases with follow-up more than 10 years in cohort-based nested case–control studies. Non-European descent.	4875:5291 (d) 4873:5299 (v)	47.3:44.6 (d) 45.7:45.8 (v)	Males 67.8 (9.7):68.8 (9.7) Females 68.0 (10.0):69.7 (8.8) (d) Males 68.2 (9.8):68.6 (9.8) Females 68.8 (9.8):69.4 (8.8) (v)
Jo 2012a and b	Korea	2004-2007 (d and v)	Hospital admission data for first admission event for colorectal cancer in individuals <55 years	Recruited from Korean Metabolic Syndrome Research Initiative study in which volunteers were offered health examinations	Low genotyping call rates (<95%), one person in each pair of biological relatives, gender mismatches, missing measurements	54:422 (a) 133:554 (b)	0:0 (a) 100:100 (b)	49.4 (7.2): 41.0 (8.0) (a) 49.4 (7.7): 42.1 (8.6) (b)
Jung 2015	South Korea	2004-2011 (d)	New cases of CRC in a cohort who had routine health assessments at health promotion centres in Seoul and GyeongGi province	Individuals within cohort of those attending routine health assessments at health promotion centres in Seoul and GyeongGi province who did not develop CRC	Prevalent CRC. Missing values for BMI, fasting blood glucose, total cholesterol, systolic blood pressure, smoking status, alcohol drinking, and exercise.	258: 145584 Sub-cohort: 173:1514	75.9:62.1 Sub-cohort: 75.0:62.4	50.7 (10.5): 41.1 (10.3) Sub-cohort: 49.7 (10.9): 40.1 (9.4)
Jung 2019	USA	1993-2014	New cases of CRC, according to centralized review of medical charts, in a cohort of postmenopausal women age 50-79 of non-Hispanic white origin recruited from 40 clinical centres nationwide and expected to live near the clinical centres for at least 3 years after enrolment.	Individuals within cohort who did not develop CRC	Follow up <1 year. Any cancer. Diabetes. Duplicates/relatives. Information on covariates unavailable.	472:6067	Not reported	66:68
Li 2015	China	2009-2012 (d)	Histological confirmed CRC	Randomly selected from patients attending outpatient clinic for health	Secondary CRC or any other cancer	1066:3880	Not reported	Not reported

Study	Country	Dates	Cases	Controls	Exclusion	N (cases:controls)	Sex	Age
				check-up with no history of any cancers or related symptoms				
Procopciuc 2017	Romania	Not given (d)	Age 30-76 with histologically confirmed CRC at a single hospital.	Normal colonoscopy with no evidence of tumours. Selected at same hospital.	FAP, Crohn's, HNPCC, missing information on dietary habits.	150:162	46.7:38.3	Males 64.7 (9.2):64.7 (8.3) Females 7.0 (10.8):63.5 (11.7)
Shiao 2018	USA	2013-2019 (d and v)	Cases age 18-80 and with life expectancy of at least 6 months in database and additional referrals by participants.	Paired family/friend members	Not given	53:53	47:26	61 (11):47 (17)
Smith 2018a and b	UK	2006-2015 (v)	New cases of CRC, as recorded in population registries, in participants in UK Biobank cohort age 40-70 and of white ethnicity	Cohort participants not diagnosed with colorectal cancer but fulfilling other inclusion criteria	Participants with missing covariate data. Related participants.	1623:359920	45% cases male	Cases: 57
Smith 2018a and c	UK	2006-2015 (v)	New cases of CRC, as recorded in population registries, in participants in UK Biobank cohort age 40-70 and of white ethnicity	Cohort participants not diagnosed with colorectal cancer but fulfilling other inclusion criteria	Participants with missing covariate data. Related participants.	1294:285583	46% cases male	Cases: 57
Wang 2013	Taiwan	Not given (d and v)	Pathologically proven colorectal cancer at one hospital	Individuals without colorectal cancer at the same hospital (no further details given)	Samples with a call rate lower than 0.97	218:385	Not reported	Not reported
Weigl 2018	Germany	2005-2013	Participants age 50-79 of German screening colonoscopy programme selected by gastroenterology practice in Southern Germany with colonoscopy reporting most advanced finding as CRC, advanced adenoma or non-advanced adenoma	Participants with colonoscopy reporting no neoplastic lesions	Questionnaire and/or genomic data not available	294:500	65:62	Not reported
Xin 2018	China	Not given (d and v)	1. Real data study: Not stated 2. Simulation study: Simulation population derived from reference dataset (Ontario Familial Colon Cancer Registry)	Not stated	Non-European descent	Simulation study: 594:665 (d) 200:200 (v) Real data study: 1316:2229	Simulation study: 43.3:54.6 Real data study: 60.3:57.9	Simulation study: 60.7 (9.2):63.9 (9.1) Real data study: 58.3 (12.9):57.9 (21.8)
Yarnall 2013	UK	N/A (v)	Simulated population of 250,000 individuals	Simulation population of 250,000 individuals	N/A	N/A	N/A	N/A

d – development, v - validation