New onset of type 2 diabetes after colorectal cancer diagnosis: Results from three prospective US cohort studies, systematic review, and meta-analysis

Hongmei Zeng,[a,b] Chen Yuan,[c] Jakub Morze,[b,d] Ruiying Fu,[a] Kai Wang,[e] Liang Wang,[e,f] Feng Sun,[g] John S. Ji,[h] Edward L. Giovannucci,[b,e,i,k,**] and Mingyang Song[b,e,j,k,*]

[a]National Central Cancer Registry, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China
[b]Department of Nutrition, Harvard School of Public Health, Boston, MA 02115, USA
[c]Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA 02115, USA
[d]Department of Cardiology and Internal Medicine, School of Medicine, University of Warmia and Mazury in Olsztyn, Olsztyn 10-082, Poland
[e]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA 02115, USA
[f]Center of Gastrointestinal Surgery, the First Affiliated Hospital, Sun Yat-sen University, Guangzhou 528406, China
[g]Department of Epidemiology and Biostatistics, School of Public Health, Peking University, 38 Xueyuan Road, Haidian District, Beijing 100191, China
[h]Vanke School of Public Health, Tsinghua University, Beijing 100191, China
[i]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115, USA
[j]Clinical and Translational Epidemiology Unit and Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02115, USA

eBioMedicine
2022;86: 104345
Published Online xxx
https://doi.org/10.1016/j.ebiom.2022.104345

Summary

Background Limited data indicate that patients with colorectal cancer (CRC) are at higher risk of developing type 2 diabetes (T2D). We prospectively examined the risk of T2D between individuals with and without CRC in three large cohorts and conducted a meta-analysis.

Methods We assessed the diagnosis of CRC and T2D among 111,485 women from the Nurses' Health Study, 112,958 women from the Nurses' Health Study II, and 46,581 men from the Health Professionals Follow-up Study. We used multivariable Cox regression with time-varying covariates to calculate the hazard ratio (HR) of T2D in relation to CRC diagnosis. We further performed a systematic review and meta-analysis of cohort studies.

Findings Up to 36 years of follow-up (6.9 million person-years), we documented 3402 incident CRC cases and 26,469 T2D cases. Compared to non-CRC individuals, those with CRC were more likely to develop T2D (multivariable-adjusted HR 1.20, 95% CI 1.05–1.38). The association was most evident for individuals with fewer risk factors for T2D. In the meta-analysis of seven cohort studies (1,061,744 participants), CRC was associated with higher T2D risk (meta-analysis HR 1.21, 95% CI 1.11–1.31, I^2 = 57.9%). By CRC duration, a statistically significant association was observed in the first 10 years but not after 10 years of CRC diagnosis (≤5 years, meta-analysis HR 1.32, 95% CI 1.27–1.36; 5.1–10 years, 1.14 [1.04–1.25]; >10 years, 1.14 [0.91–1.37]).

Interpretation CRC was associated with increased T2D risk, especially in the first ten years after CRC diagnosis. Our findings highlight the importance of T2D prevention for CRC survivorship care.

Funding NHS cohort infrastructure grant (UM1 CA186107), NHS program project grant that funds cancer research (P01 CA87969), NHS II cohort infrastructure grant (U01 CA176726), HPFS cohort infrastructure grant (U01 CA167552) and the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2021-I2M-1-010).

Keywords: Type 2 diabetes mellitus; Colorectal neoplasms; Prospective study; Meta analysis

*Corresponding author. 677 Huntington Avenue, Boston, MA 02115, USA.
**Corresponding author. 677 Huntington Avenue, Boston, MA 02115, USA.
E-mail addresses: msong@hsph.harvard.edu (M. Song), egiovann@hsph.harvard.edu (E.L. Giovannucci).
[k]Equal contribution as senior authors and corresponding authors.

Research in context

Evidence before the study

We searched PubMed, Embase, and Web of Science for prospective studies to examine the relationship between colorectal cancer (CRC) and subsequent risk of type 2 diabetes (T2D), without language or date restrictions. We screened the reference lists of selected studies to identify additional relevant studies. We found four prospective studies that have examined the relationship between CRC and risk of T2D. The findings from the previous studies were inconclusive. Given that both diseases share common risk factors and that T2D is an established risk factor of CRC, it is important to simultaneously address multiple methodological issues in the association between CRC and subsequent T2D risk, including reverse causality, detection bias, and influence of shared risk factors. However, the association has not been addressed in a large prospective cohort with long-term follow-up, a large sample size, and accounting for time-varying confounding. Furthermore, meta-analyses that explored the association between CRC and T2D risk overall and by CRC duration were lacking.

Added value of the study

To our knowledge, this study significantly exceeds the scale of all previous research by combining the most comprehensive updated systematic review of the topic with new primary analysis from three large prospective cohorts with serial-collected information on lifestyle and detailed disease diagnosis. Results from the three cohorts and from the meta-analysis demonstrate that, after accounting for shared risk factors, the onset of T2D in CRC patients may be at least partly due to CRC itself or CRC-related treatment, especially in the first ten years after CRC diagnosis. Furthermore, we found that the association was most evident for individuals with fewer risk factors for T2D.

Implications of all available evidence

As the number of CRC survivors continues to increase due to a growing and aging population, evidence-based cancer survivorship care is vital to improve patients' prognosis and quality of life. Our findings have important public health implications. Our study results highlight the urgent need for improved recognition and diagnosis of T2D associated with CRC. It is important to establish practice guidelines for the prevention, diagnosis, and treatment of CRC-related T2D. Furthermore, enhanced T2D prevention for CRC survivors are needed among all patients, regardless of risk profiles.

Introduction

Colorectal cancer (CRC) is the fourth most common cancer in the United States (US) and worldwide.[1] As the number of CRC survivors continues to increase due to a growing and aging population, evidence-based cancer survivorship care is vital to improve patients' prognosis and quality of life.[2] A limited body of evidence suggests that CRC is associated with an increased risk of type 2 diabetes mellitus (T2D).[3–5] However, the evidence of a positive association remains inconclusive.[6] Although both diseases share common risk factors and T2D is an established risk factor of CRC,[7] few previous studies simultaneously addressed multiple methodological challenges in the association between CRC and subsequent T2D risk, including reverse causality (T2D as a risk factor of CRC),[8] detection bias, and influence of shared risk factors (unhealthy diet, obesity, etc.).[7] Therefore, it is important to quantify the subsequent risk of T2D incidence in CRC patients, adjusting for potential confounding factors and addressing reverse causation.

The aim of our study was to prospectively evaluate the association between CRC diagnosis and T2D incidence using a prospective cohort design with a long-term follow-up, large sample size and detailed data on confounders and effect modifiers. We also quantified the T2D risk according to CRC duration and various lifestyle and clinical characteristics. We leveraged data from three large ongoing prospective cohort studies: the Nurses' Health Study (NHS), the NHS II, and the Health Professionals Follow-up Study (HPFS). In these three cohorts, we collected detailed information on lifestyle factors and health conditions every two years and diet every four years. Furthermore, we conducted an updated systematic review of the literature and meta-analysis of cohort studies that included these new results, aiming to provide a comprehensive overview of the existing evidence.

Methods

Study design and data resources

Study population

In this prospective cohort study, we used data from three ongoing cohorts of medical professionals in the US. The NHS started in 1976, and recruited 121,700 female registered nurses aged 30–55 in 11 US states.[9] The NHS II began in 1989 and composed 116,429 female nurses aged 25–42 in 14 US states. The HPFS was initiated in 1986, and enrolled 51,529 male health professionals aged 40–75 years in 50 US states. We mailed follow-up questionnaires biennially to the cohort participants and collected their updated information on lifestyle factors, medication use, and disease diagnosis.

The cumulative follow-up rate exceeded 90% in all three cohorts. We used 1980 for the NHS, 1989 for the NHS II, and 1986 for the HPFS as the baseline, when we first collected detailed lifestyle data. We excluded participants who reported type 1 diabetes, T2D or cancer (except nonmelanoma skin cancer) at baseline, who had completed a baseline questionnaire only, stage IV/non-first primary CRC cases, and those diagnosed with T2D and CRC in the same month. The exclusion criteria were based on considerations of minimizing detection bias due to reverse causality and competing risk of death. After exclusion, the final analyses included 111,485 women from the NHS, 112,958 women from the NHS II, and 46,581 men from the HPFS (Appendix p 9).

The study was approved by the Institutional Review Board at the Brigham and Women's Hospital and the Harvard T.H. Chan School of Public Health, and those participating registries as required. Informed consent was implied by completion of the questionnaire for the entire cohort. Written informed consent for medical record ascertainment and review was obtained from CRC patients.

Assessment of CRC cases

The self-reported diagnosis of CRC was based on a biennial follow-up questionnaire. The study physicians who were blinded to the exposure information ascertained the CRC diagnosis by review of medical records and pathologic reports. The disease stage, histologic findings, and tumour location were also recorded during the medical records review. For nonrespondents, the National Death Index was used to ascertain any diagnosis of CRC that contributed to death, and the study physicians further asked permission from the next of kin of decedents to obtain medical records. We used the American Joint Committee on Cancer (AJCC) tumour–node–metastasis (TNM) cancer staging system to classify CRC stage at diagnosis.[10]

Assessment of T2D cases and death

The primary outcome of the current study was incident T2D. When participants reported T2D diagnosis in a biennial questionnaire, we sent a validated supplementary questionnaire regarding diagnostic tests, symptoms, and hypoglycemic therapy. Before 1998, the National Diabetes Data Group criteria were used to diagnose diabetes mellitus: (1) elevated glucose concentration (fasting glucose concentration ≥7.8 mmoL/L, random blood glucose ≥11.1 mmoL/L or plasma glucose ≥11.1 mmol/L 2 h after an oral glucose load), together with one or more diabetes mellitus-related symptoms (weight loss, polyuria, excessive thirst or hunger); (2) no symptoms but raised concentration of glucose on at least of two occasions; or (3) treatment with insulin or other hypoglycemic drugs. For participants with T2D identified after 1998, the cut-off point for fasting plasma concentrations of glucose was lowered to 7.0 mmoL/L, according to the criteria of the American Diabetes Association. After 2010, we further included hemoglobin A1c ≥6.5% in the criteria.[11] In the validation study of T2D confirmation, more than 97% of self-reported T2D cases were confirmed by medical record review in the cohorts.[12,13]

Deaths were identified by reports from next of kin, the US postal service, state vital statistics department, and systematic searches of the National Death Index. Using these methods, we were able to ascertain more than 98% of deaths in each cohort.[14]

Assessment of covariates and effect modifiers

We collected detailed information on participants' race, height, family history of diabetes, and myocardial infarction. We sent follow-up questionnaires every two years to collect and update lifestyle data, and potential risk factors, including smoking status, use of vitamin supplements, weight, physician-diagnosed hypertension, hypercholesterolemia, aspirin, and multivitamin use. Body mass index (BMI) was calculated as weight in kilograms divided by the square of height in meter (kg/m^2). We used food frequency questionnaires every four years to obtain information on usual intake of food and beverages. Physical activity was repeatedly assessed in the three cohorts every two or four years and was calculated in average metabolic equivalent task (MET) –h/wk.[15] We further calculated the overall dietary score according to the 2010 Alternate Healthy Eating Index (AHEI) to measure the diet quality.[16] The validity and reproducibility of these self-reported covariates have been reported previously.[15,17,18]

Statistical analysis

We presented the demographic, lifestyle, and dietary information according to the person-time of CRC status, to capture the dynamic characteristics of participants with and without CRC during follow-up. We calculated person-years of follow-up from the date of return of the baseline questionnaire until the date of T2D diagnosis, death, or the end of follow-up (June 2016 in NHS, June 2015 in NHS II, January 2016 in HPFS), whichever came first. For the primary analysis, CRC status was treated as a time-varying exposure. For participants diagnosed with CRC as the first cancer diagnosis, the follow-up time before CRC diagnosis was counted in the non-exposure group, and the follow-up time after CRC diagnosis was counted in the exposure group. We censored those who were lost to follow-up at the return date of the last questionnaire. We also censored those who had a cancer diagnosis other than CRC after cancer diagnosis based on the consideration that cancer diagnosis other than CRC may be also related to subsequent risk of T2D. CRC duration was defined in years as the difference between the questionnaire return date in

each follow-up cycle and the date of CRC diagnosis. Analyses were first conducted within each cohort separately, and then in the pooling data from the three cohorts in the absence of heterogeneity of results.

We used multivariable Cox regression with updated time-varying covariates to examine the association between CRC and T2D. The proportion hazards assumption was evaluated by including an interaction term between CRC status and the duration of follow-up. No evidence of violations of the assumption was detected. All analyses were stratified by cohort, age, and follow-up intervals. We used three models with increasing degrees of adjustment to account for potential confounding. Model 1 was adjusted for age. We used model 2 as the main multivariable model in our study, adjusting for race (white, other), family history of myocardial infarction (no, yes), smoking status (never, former, current), body mass index as a continuous variable, alcohol intake (g/day in quintiles), physical activity MET (quintiles), multivitamin use (no, yes), aspirin use (no, yes), AHEI-2010 (quintiles), and family history of diabetes (no, yes). In model 3, we further adjusted for status of hypertension and hypercholesterolemia, which have potentially overlapping causal pathways for CRC and T2D, to examine if any association between CRC and T2D were independent of these. We calculated the hazard ratios (HRs) according to categories of T2D duration using prespecified cutoffs: ≤5 years, 5.1–10 years and >10 years. When data on lifestyle factors were missing in each questionnaire cycle, we carried forward values from the last non-missing observation. We also assessed the associations by anatomical site and AJCC TNM stage of CRC using the subtype-stratified Cox proportional hazards regression model by the duplication method.[19]

To evaluate potential effect modification, we performed stratified analyses according to age, sex, family history of diabetes, hypertension, hypercholesterolemia, BMI status, AHEI score, smoking status, physical activity MET-h/wk, aspirin use, family history of CRC, and alcohol intake. We further constructed a T2D risk score to identify subgroups with different levels of predisposing risk (Appendix p 2).[20,21] The T2D risk score (range 0–6) was based on information on the following six risk factors—family history, BMI, smoking, physical activity, alcohol assumption and diet, given that they were established risk factors of T2D and included in previously validated score.[21–23] We added the interaction terms to the models to calculate the P value for interaction using Wald test.

To assess the robustness of primary findings, we performed several sensitivity analyses. First, we excluded T2D cases that occurred within six months and within the first year after CRC to minimise detection bias and reverse causality. Second, we excluded CRC cases who died within two years after diagnosis and repeated the analyses to examine whether our analyses were affected by severity of the disease. Third, given that patients may lose weight due to CRC diagnosis, we used the subjects' lifetime maximum BMI status instead of the updated BMI in the multivariable models. Fourth, we combined the three cohorts' estimates using a fixed-effect meta-analysis given that results did not show heterogeneity among cohorts. Statistical analyses were performed with SAS software, version 9.4 (SAS Institute, Cary, North Carolina). Two-sided $P < 0.05$ was considered statistically significant.

Systematic review and meta-analysis

We performed a systematic review and meta-analysis including the current study and previous prospective studies that evaluated the association between CRC and risk of T2D. Two authors (R.F. and J.M.) independently screened the literature, extracted the data, and conducted the risk of bias assessment in each study. Any discrepancies between the investigators were resolved through discussion and consensus with the corresponding authors. We searched PubMed, Embase, and Web of Science during 11th May 2021 and 16th May 2021, without language or date restrictions (Fig. 1). We presented adopted search strategies (Appendix p 6). Moreover, we screened the reference lists of retrieved reports to identify additional relevant studies. We included studies with: (1) prospective observational design (cohort, case-cohort, nested case–control); (2) conducted in adults; and (3) which reported estimates for the association between CRC and incidence of T2D or if they provided risk estimates for duration of CRC. Extracted data included first author name, year of publication, study design, country, number of participants, mean age of the study population, follow-up period, number of CRC cases and T2D cases, adjusted covariates, as well as the risk estimates (95% CIs) from the multivariable model overall and by CRC duration. For studies reporting several risk estimates, we selected the one with maximal adjustment. One study only reported HRs and 99% CIs, and we treated the 95% CIs the same as 99% CIs for a conservative estimate.[5] When studies reported risk estimates separately within a study, risk estimates were pooled by using a fixed-effect meta-analysis.

The methodological quality of each study was assessed according to the Newcastle–Ottawa Quality Assessment Scale (NOS).[24] Age, sex and BMI were considered as primary confounders of the association between CRC and T2D risk. Smoking, drinking, comorbidity score, and hypertension status were considered as secondary confounders.

We fitted restricted maximum likelihood (REML) random effects models to compute the pooled relative risk for T2D. Statistical heterogeneity between studies was evaluated with Cochrane Q test and I^2 statistic, with I^2 of 50% indicating substantial heterogeneity. We used

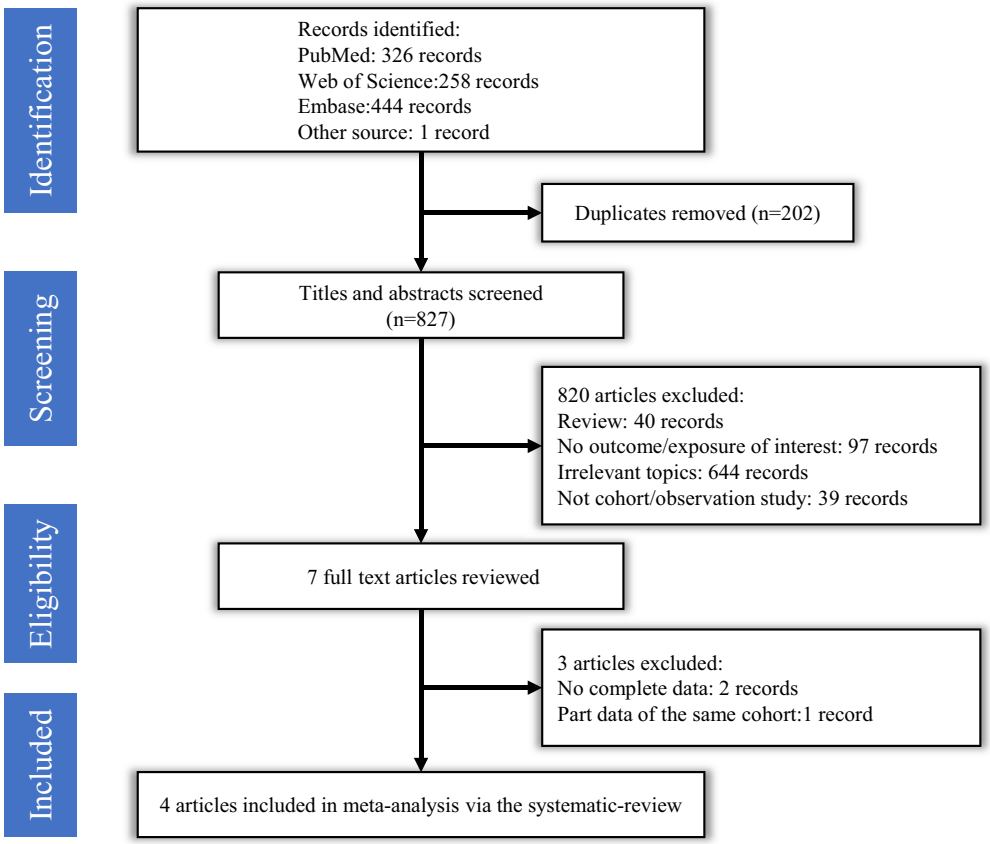

Fig. 1: Flowchart of the study selection for systematic review and meta-analysis.

fixed-effect models to compute the pooled relative risk between CRC duration and T2D given that there was no prior knowledge indicating a sex difference. Publication bias was assessed using funnel plot, the Begg's and Egger's statistical test. We conducted an influence study by systematically removing each study from the meta-analysis to evaluate if any single study may cause heterogeneity. We conducted meta-analysis using STATA/MP version 16 software (Stata Corp, College Station, Texas, USA).

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation or writing of the report. The corresponding authors had full access to all the data and had responsibility for the decision for paper submission and publication.

Results

Cohort analyses

Our study included a total of 271,024 participants. The NHS and the HPFS cohorts were older, had a higher percentage of population who were smoking, drinking, with aspirin use, with no family history of T2D, and with hypertension than those from the NHS II cohort. During 6,893,477 person-years of follow-up (3,176,391 person-years in NHS, 2,701,440 person-years in NHS II, 1,015,646 person-years in HPFS), we documented 3402 incident CRC cases (2079 in NHS, 405 in NHS II, 918 in HPFS) and 26,469 incident T2D cases (13,833 in NHS, 8686 in NHS II, 3950 in HPFS). Among the CRC cases, 209 developed T2D after their cancer diagnosis. The age-standardised basic characteristics of study participants according to CRC status are shown in Table 1. Compared to non-CRC participants, those who developed CRC were older and more likely to be overweight, have hypertension, and hypercholesterolemia.

In the absence of heterogeneity across three cohorts, we performed pooled analysis of individual-participant data. The crude incidence of T2D among CRC patients was 578 per 100,000, which was higher than among those without CRC (383 per 100,000). After adjustment for age, participants who had CRC had a significantly higher risk for T2D compared with non-CRC participants (model 1-HR, 1.26, 95% CI 1.10–1.45). After further accounting for updated lifestyle

	All	NHS		All	NHS II		All	HPFS	
		No CRC	CRC[a]		No CRC	CRC[a]		No CRC	CRC[a]
Age[b]	60.6 (11.6)	60.5 (11.5)	71.6 (9.3)	46.2 (8.7)	46.2 (8.7)	54.3 (6.5)	63.7 (11.2)	63.6 (11.2)	74.2 (9.5)
Age group[b]									
<60 y, %	48.7	49.0	11.6	93.9	93.9	78.5	38.0	38.3	7.8
60–64 y, %	15.5	15.6	12.2	5.4	5.3	18.6	16.3	16.4	9.3
65–69 y, %	13.7	13.7	17.1	0.7	0.7	2.9	16.4	16.4	15.6
70–74 y, %	10.3	10.2	20.5	0.0	0.0	0.0	12.4	12.4	18.6
≥75 y, %	11.8	11.6	38.7	0.0	0.0	0.0	16.8	16.5	48.7
White, %	96.9	96.9	97.1	92.7	92.7	95.9	95.0	95.0	96.8
Body mass index, kg/m^2	25.4 (4.7)	25.4 (4.7)	26.0 (5.1)	26 (5.8)	26.0 (5.8)	26.5 (5.6)	25.7 (3.5)	25.7 (3.5)	26.0 (3.4)
Alternative healthy eating index	53.1 (11.9)	53.1 (11.9)	52.8 (11.1)	53.3 (13.3)	53.3 (13.3)	56.5 (13.8)	56 (12.3)	56.0 (12.3)	56.0 (11.7)
Physical activity, MET-h/wk	17.2 (22.8)	17.2 (22.8)	15.9 (21.8)	22.3 (29.9)	22.3 (29.9)	20.9 (29.9)	31.5 (31.6)	31.5 (31.6)	30.8 (31.8)
Alcohol consumption, g/day	6 (10.4)	6 (10.4)	5.6 (9.6)	4.2 (7.9)	4.2 (7.9)	4.4 (8.5)	11.4 (15.1)	11.4 (15.1)	14 (16.7)
Current drinking, %	47.3	47.3	48.1	60.8	60.8	58.5	75.1	75.0	78.7
Never smoking, %	47.6	47.7	46.8	65.9	65.9	61.2	55.8	55.9	51.6
Family history of diabetes, %	24.4	24.3	26.5	37.0	37.0	37.4	19.9	19.9	23.3
Family history of myocardial infarction, %	23.7	23.7	22.7	21.3	21.3	21.2	15.7	15.7	12.3
Multivitamin use, %	41.8	41.8	45.6	45.1	45.1	46.9	45.6	45.6	51.1
Aspirin use, %	43.3	43.3	40.5	17.9	17.9	24.9	40.0	40.0	43.2
Hypercholesterolemia status, %	29.1	29.1	32.5	21.1	21.1	27.6	31.0	30.9	40.5
Hypertension status, %	27.4	27.3	33.7	14.3	14.3	19.0	25.5	25.5	28.7

Abbreviations: T2D, type 2 diabetes; HPFS, Health Professionals Follow-up Study; NHS, Nurses' Health Study; MET, Metabolic equivalent from recreational and leisure-time activities. Values are means (SD) for continuous variables; percentages for categorical variables and are standardized to the age distribution of the study population. Values of polytomous variables may not sum to 100% due to rounding. [a]The number of CRC cases in NHS, NHS II and HPFS cohort were 2079, 405 and 918, respectively. [b]Age and age groups were not age standardized.

Table 1: **Age-standardized basic characteristics of study participants according to colorectal cancer (CRC) status (person-years).**

factors and other confounders in the multivariable model, CRC was significantly associated with a higher T2D risk (model 2-HR 1.20, 95% CI 1.05–1.38). The association was slightly attenuated but remained statistically significant after further adjusting for hypertension and hypercholesterolemia status (Table 2). When assessed by CRC duration, we found a positive association with T2D risk for CRC duration within the first five years (model 1-HR 1.22, 95% CI 0.99–1.51) and 5.1–10 years (model 1-HR 1.35, 95% CI 1.05–1.73) in the age-adjusted model. The association between CRC and T2D risk slightly attenuated after 10 years of diagnosis compared to 5.1–10 years. In multivariable models, we observed a weaker non-significant association with T2D risk for CRC duration within five years (model 2-HR 1.16, 95% CI 0.94–1.44) and 5.1–10 years (model 2-HR 1.26, 95% CI 0.98–1.61).

For the analyses according to CRC stage and subsite (Table 3), similar positive associations were observed for early (stage I) and non-early stage (stage II/III/unknown) CRC. According to tumour subsites, the positive association appeared to be stronger among patients with rectal cancer (model 2-HR 1.57, 95% CI 1.27–1.94) than with colon cancer (model 2-HR 1.15, 95% CI 0.99–1.34). In the stratified analysis by potential risk factors for T2D (Fig. 2), the association between CRC and T2D risk was significantly stronger among patients under age 70 years, with use of aspirin, without family history of CRC and without hypertension (P for interaction<0.05). In addition, the association between CRC and T2D risk was strongest in subjects with the lowest T2D risk score (HR 1.77, 95% CI 1.35–2.33, P for interaction<0.001). In contrast, we did not find significant interaction with other factors such as smoking or BMI status.

In sensitivity analyses, we excluded CRC cases diagnosed within the first six months or within the first year and the results did not materially change (Appendix p 3). We further excluded those who died within the first two years after CRC diagnosis and found that the positive association remained robust (Appendix p 3). The associations between CRC and T2D risk remained significant when we used life-time maximum BMI instead of updated BMI status for covariate adjustment (Appendix p 4), as well as when we used a fixed-effect meta-analysis to pool the results (Appendix p 5).

Systematic review and meta-analysis

After removing the duplicates, a total of 827 studies were screened, and four cohort studies (seven including the three cohorts from the current study) met eligibility criteria (Fig. 1). The characteristics of included studies are described and evaluated in Table 4. We found that all included studies were of high quality (a NOS score of 8 or more, Appendix p 7).

Cohort	Exposure status	T2D cases/person-years	Crude incidence rate per 100,000	Model 1 HR (95% CI)	Model 2 HR (95% CI)	Model 3 HR (95% CI)
NHS	**CRC**					
	No	13,691/3,153,304	434	1 (Referent)	1 (Referent)	1 (Referent)
	Yes	142/23,087	615	1.27 (1.07, 1.50)	1.19 (1.01, 1.40)	1.17 (0.99, 1.38)
	CRC duration					
	No CRC	13,691/3,153,304	434	1 (Referent)	1 (Referent)	1 (Referent)
	≤5 y	54/8882	608	1.18 (0.91, 1.55)	1.10 (0.84, 1.44)	1.08 (0.83, 1.42)
	5.1–10 y	41/5943	689	1.35 (0.99, 1.83)	1.23 (0.90, 1.67)	1.21 (0.89, 1.64)
	>10 y	47/8263	569	1.31 (0.98, 1.75)	1.27 (0.95, 1.69)	1.24 (0.93, 1.65)
NHS II	**CRC**					
	No	8662/2,698,045	321	1 (Referent)	1 (Referent)	1 (Referent)
	Yes	24/3396	707	1.67 (1.12, 2.49)	1.74 (1.16, 2.60)	1.74 (1.16, 2.59)
	CRC duration					
	No CRC	8662/2,698,045	321	1 (Referent)	1 (Referent)	1 (Referent)
	≤5 y	10/1710	585	1.44 (0.77, 2.67)	1.45 (0.78, 2.69)	1.49 (0.80, 2.77)
	5.1–10 y	8/966	828	1.92 (0.96, 3.84)	2.03 (1.02, 4.06)	2.06 (1.03, 4.11)
	>10 y	6/721	832	1.85 (0.83, 4.12)	2.03 (0.91, 4.51)	1.87 (0.84, 4.17)
HPFS	**CRC**					
	No	3907/1,005,942	388	1 (Referent)	1 (Referent)	1 (Referent)
	Yes	43/9703	443	1.10 (0.81, 1.48)	1.04 (0.77, 1.41)	1.02 (0.75, 1.38)
	CRC duration					
	No CRC	3907/1,005,942	388	1 (Referent)	1 (Referent)	1 (Referent)
	≤5 y	22/3966	555	1.24 (0.82, 1.89)	1.18 (0.77, 1.79)	1.14 (0.75, 1.73)
	5.1–10 y	13/2627	495	1.15 (0.67, 1.98)	1.06 (0.62, 1.83)	1.05 (0.61, 1.81)
	>10 y	8/3111	257	0.79 (0.39, 1.58)	0.77 (0.39, 1.55)	0.76 (0.38, 1.52)
Pooled cohorts[a]	**CRC**					
	No	26,260/6,857,291	383	1 (Referent)	1 (Referent)	1 (Referent)
	Yes	209/36,187	578	1.26 (1.10, 1.45)	1.20 (1.05, 1.38)	1.18 (1.03, 1.35)
	CRC duration					
	No CRC	26,260/6,857,291	383	1 (Referent)	1 (Referent)	1 (Referent)
	≤5 y	86/14,558	591	1.22 (0.99, 1.51)	1.16 (0.94, 1.44)	1.14 (0.92, 1.41)
	5.1–10 y	62/9535	650	1.35 (1.05, 1.73)	1.26 (0.98, 1.61)	1.24 (0.97, 1.59)
	>10 y	61/12,094	504	1.24 (0.96, 1.59)	1.21 (0.94, 1.56)	1.18 (0.92, 1.52)

Abbreviations: CI, confidence interval; HR, hazard ratio. Multivariable adjusted HRs were estimated from Cox proportional regression model. Model 1: stratified by age (<60, 60–64, 65–69, 70–74, ≥75 years), cohort and follow-up cycle (2-year interval). Model 2 (main multivariable model): model 1 + race (white, other), family history of myocardial infarction (no, yes), smoking status (never, former, current), body mass index as a continuous variable, alcohol intake (0, 0.1–4.9,5–14.9,15–29.9, ≥30 g/day), physical activity MET (quintiles), multivitamin use (no, yes), aspirin use (no, yes), AHEI-2010 (quintiles) and family history of diabetes (no, yes). Model 3: Model 2 + hypertension (no, yes) and hypercholesterolemia (no, yes). [a]P for heterogeneity across cohorts = 0.262.

Table 2: **Association between colorectal cancer (CRC) diagnosis and risk of type 2 diabetes (T2D).**

In the meta-analysis of seven prospective cohort studies (1,061,744 participants), CRC was associated with T2D risk (Fig. 3A, meta-analysis HR 1.21, 95% CI 1.11–1.31, I^2 = 57.9%, P = 0.009). In meta-analysis stratified by CRC duration (Fig. 3B), the association was statistically significant in the first ten years after CRC diagnosis (≤5 years, meta-analysis HR 1.32, 95% CI 1.27–1.36, I^2 = 43.6%; 5.1–10 years, meta-analysis HR 1.14, 95% CI 1.04–1.25, I^2 = 0.0; >10 years, meta-analysis HR 1.14, 95% CI 0.91–1.37, I^2 = 0.0). We found no statistical evidence for publication bias (P = 0.58 for Egger's test and 0.76 for Begg's test, Appendix p 13). Sensitivity analysis showed that exclusion of one study at a time yielded relatively stable risk estimates (Appendix p 8).

Discussion

Leveraging data from three large prospective cohorts, we found that, after adjusting for shared risk factors, patients with CRC were more likely to develop incident T2D compared with those without CRC. Results from the meta-analysis lend further support to the overall positive association between CRC and subsequent T2D risk. With careful adjustment of important confounders in all analyses and explicitly addressing potential reverse causality, the positive association between CRC and T2D risk remained statistically significant. Furthermore, the association was primarily observed among individuals under age 70 years, without hypertension, and with the lowest T2D risk score. Our meta-analysis provides compelling evidence for a positive association between

Group	T2D cases/person-years	Crude incidence rate per 100,000	Model 1 HR (95% CI)	Model 2 HR (95% CI)	Model 3 HR (95% CI)
No CRC	26,260/6,857,291	383	1 (Referent)	1 (Referent)	1 (Referent)
By stage[a]					
Stage I CRC	60/10,522	570	1.26 (0.98, 1.62)	1.24 (0.96, 1.59)	1.16 (0.90, 1.49)
Stage II/III/unknown CRC	149/25,664	581	1.26 (1.08, 1.49)	1.19 (1.01, 1.40)	1.19 (1.01, 1.40)
By subsite[a]					
Colon cancer	166/29,772	560	1.22 (1.04, 1.42)	1.15 (0.99, 1.34)	1.13 (0.97, 1.31)
Rectal cancer	85/11,300	752	1.67 (1.35, 2.07)	1.57 (1.27, 1.94)	1.56 (1.26, 1.93)

Abbreviations: CI, confidence interval; HR, hazard ratio. Multivariable adjusted HRs were estimated from Cox proportional regression model. Model 1: stratified by age (<60, 60–64, 65–69, 70–74, ≥75 years), cohort and follow-up cycle (2-year interval). Model 2 (main model): model 1 + race (white, other), family history of myocardial infarction (no, yes), smoking status (never, former, current), body mass index as a continuous variable, alcohol intake (0, 0.1–4.9, 5–14.9, 15–29.9, ≥30 g/day), physical activity MET (quintiles), multivitamin use (no, yes), aspirin use (no, yes), AHEI-2010 (quintiles) and family history of diabetes (no, yes). Model 3: Model 2 + hypertension (no, yes) and hypercholesterolemia (no, yes). [a]P for heterogeneity across cohorts were assessed using meta-analysis.

Table 3: **Association of CRC and T2D by cancer subsite and stage.**

CRC and risk of T2D especially in the first ten years after CRC diagnosis. The study results highlight the importance of T2D prevention for CRC survivorship care.

The findings for the positive association between CRC and T2D development are in accordance with most existing studies.[3–5] Our study extended these data in several important ways. First, the three cohorts we used in our study are well established prospective cohorts that had a large sample size and high follow-up rate for more than 30 years, which allows us to capture a more comprehensive picture of the relationship between CRC and T2D. We account for multiple factors that may influence the relationship between CRC and T2D, including CRC subtype, stage, and CRC duration. Second, the repeated measurement of lifestyle variables and disease status confirmed by comprehensive medical review allows us to conduct time-varying analysis adjusting for potential confounders, and to minimise reverse causality and detection bias. By using prospectively ascertained and updated covariate information, our analysis provides a robust estimate for the magnitude of the association between CRC diagnosis and incident T2D. Third, the meta-analysis confirms our cohort findings and provides strong evidence that T2D prevention should be an important public health imperative in CRC survivors.

Our study results further demonstrate that, after accounting for shared risk factors, the new onset of T2D after CRC diagnosis may be at least partly due to CRC itself or CRC-related treatment. The independent risk of T2D due to CRC was supported by the following findings of our study. First, even after adjusting for some shared risk factors for both diseases such as obesity, physical inactivity and hypertension, the association remains statistically significant. Second, we observed positive findings in subjects of stage I cases and CRC survivors after five years who do not generally receive chemotherapy. Supporting our results, a previous study demonstrated that individuals were at a high risk of developing T2D after diagnosis of CRC irrespective of chemotherapy.[5] Third, our stratified results showed that the association was more pronounced in CRC patients who are younger, without hypertension, without family history of CRC, and having lower predisposing T2D risk score, suggesting that CRC and/or related treatment may increase T2D risk preferentially in those with fewer traditional risk factors for T2D risk. These results may indicate the excess diabetes cases that occur after a diagnosis of CRC may have features that may be etiologically distinct from typical T2D. Fourth, we did not find sex differences in the association of CRC and T2D, consistent with previous studies.[4,5] However, the positive association between CRC and T2D appeared to be stronger among patients with rectal cancer.[25] We study adds further evidence for site-specific differences of CRC. More detailed research with larger sample size and the potential mechanism are desirable to examine the subsite-differences of CRC on T2D risk.

The mechanisms underlying the independent influence on T2D by CRC or CRC-related treatment are not established. Gut microbiota alteration and inflammation associated with CRC may lead to increased risk of T2D.[25–27] Dysbiosis of gut microbiota, inflammation and metabolic disturbances such as insulin resistance and dysregulated metabolic pathways frequently occur in CRC patients,[28–30] which may contribute to the development of T2D after CRC.[31,32] Many anticancer agents may directly influence the risk of T2D and hyperglycaemia for cancer patients.[33,34] Moreover, synthetic steroids such as dexamethasone, which are widely used drugs to accompany chemotherapy use in CRC patients, may impair glucose metabolism and exacerbate insulin resistance.[35,36] Hyperglycaemia is also reported to be associated with platinum-based chemotherapy,[37,38] which may not only have a toxic effect on pancreatic β cells, but also inhibit the glycolytic enzymes, contributing to inhibition of insulin. Given the interrelated conditions and bidirectional effects between metabolic disturbance and cancer, more experimental studies are needed to further understand the mechanisms of T2D development in CRC patients.

Stratification variables	Cases/person-years (incidence rate per 100 000) No CRC	CRC	HR (95% CI)	P_{heter}
Age, years				0·002
<60	13 860/4 449 445 (311)	35/6001 (583)	1·43 (1·03, 2·00)	
60-69	7941/1 391 030 (570)	82/9659 (848)	1·37 (1·10, 1·70)	
≥70	4459/1 016 815 (439)	92/20 527 (448)	1·03 (0·84, 1·27)	
Gender				0·379
Women	3907/1 005 942 (388)	43/9703 (443)	1·25 (1·07, 1·45)	
Men	22 353/5 851 348 (382)	166/26 483 (627)	1·04 (0·77, 1·41)	
Family history of diabetes				0·086
No	13 182 /4 896 924 (269)	124/26 818 (462)	1·26 (1·06, 1·51)	
Yes	13 078/1 960 367 (667)	85/9368 (907)	1·13 (0·91, 1·39)	
Hypertension status				0·015
No	12 635/5 371 527 (235)	85/21 067 (403)	1·37 (1·11, 1·70)	
Yes	13 625/1 485 764 (917)	124/15 120 (820)	1·06 (0·89, 1·27)	
Hypercholesterolemia status				0·704
No	13 913/5 074 048 (274)	89/21 558 (413)	1·13 (0·92, 1·40)	
Yes	12 347/1 783 243 (692)	120/14 628 (820)	1·27 (1·06, 1·53)	
BMI status, kg/m^2				0·250
<25	5108/3 959 965 (129)	54/18 407 (293)	1·60 (1·22, 2·10)	
25-29·9	7491/1 870 161 (401)	53/12 034 (440)	0·81 (0·62, 1·07)	
≥30	13 661/1 027 166 (1330)	102/5746 (1775)	1·24 (1·02, 1·51)	
AHEI-2010 score				0·825
Tertile 1	9195/1 918 627 (479)	74/11 112 (666)	1·21 (0·96, 1·52)	
Tertile 2	10 699/2 876 968 (372)	82/13 677 (600)	1·22 (0·98, 1·52)	
Tertile 3	6366/2 061 695 (309)	53/11 398 (465)	1·18 (0·90, 1·55)	
Smoking status				0·130
Never	13 533/3 867 491 (350)	101/16 888 (598)	1·36 (1·12, 1·66)	
Past	9967/2 238 244 (445)	96/17 076 (562)	1·09 (0·89, 1·34)	
Current	2760/751 556 (367)	12/2223 (540)	1·06 (0·60, 1·88)	
Physical activity, MET-h/wk				0·305
<7·5	14 996/3 312 527(453)	131/19 403 (675)	1·31 (1·10, 1·56)	
≥7·5	11 264/3 544 764 (318)	78/16 783 (465)	1·07 (0·85, 1·33)	
Alcohol intake				0·663
Low	21 320/4 886 609 (436)	164/23 730 (691)	1·26 (1·08, 1·47)	
Moderate	3603/1 383 997 (260)	31/8244 (376)	0·99 (0·69, 1·41)	
Heavy	1337/586 684 (228)	14/4213 (332)	1·09 (0·64, 1·85)	
Aspirin use				0·005
No	15 069/4 585 045 (329)	129/20 582 (627)	1·31 (1·13, 1·53)	
Yes	11 272/2 272 902 (496)	83/15 601 (532)	0·94 (0·67, 1·32)	
Family history of CRC				0·046
No	19 403/4 902 545 (396)	177/28 203 (628)	1·48 (1·24, 1·76)	
Yes	6983/1 954 403(357)	35/7981 (439)	0·94 (0·76, 1·17)	
T2D risk score				<0·001
Low	5198/3 740 535 (139)	53/15 546 (341)	1·77 (1·35, 2·33)	
Mid	15 380/2 757 648 (558)	107/17 287 (619)	1·01 (0·83, 1·22)	
High	5682/359 108 (1582)	49/3354 (1461)	1·09 (0·82, 1·44)	

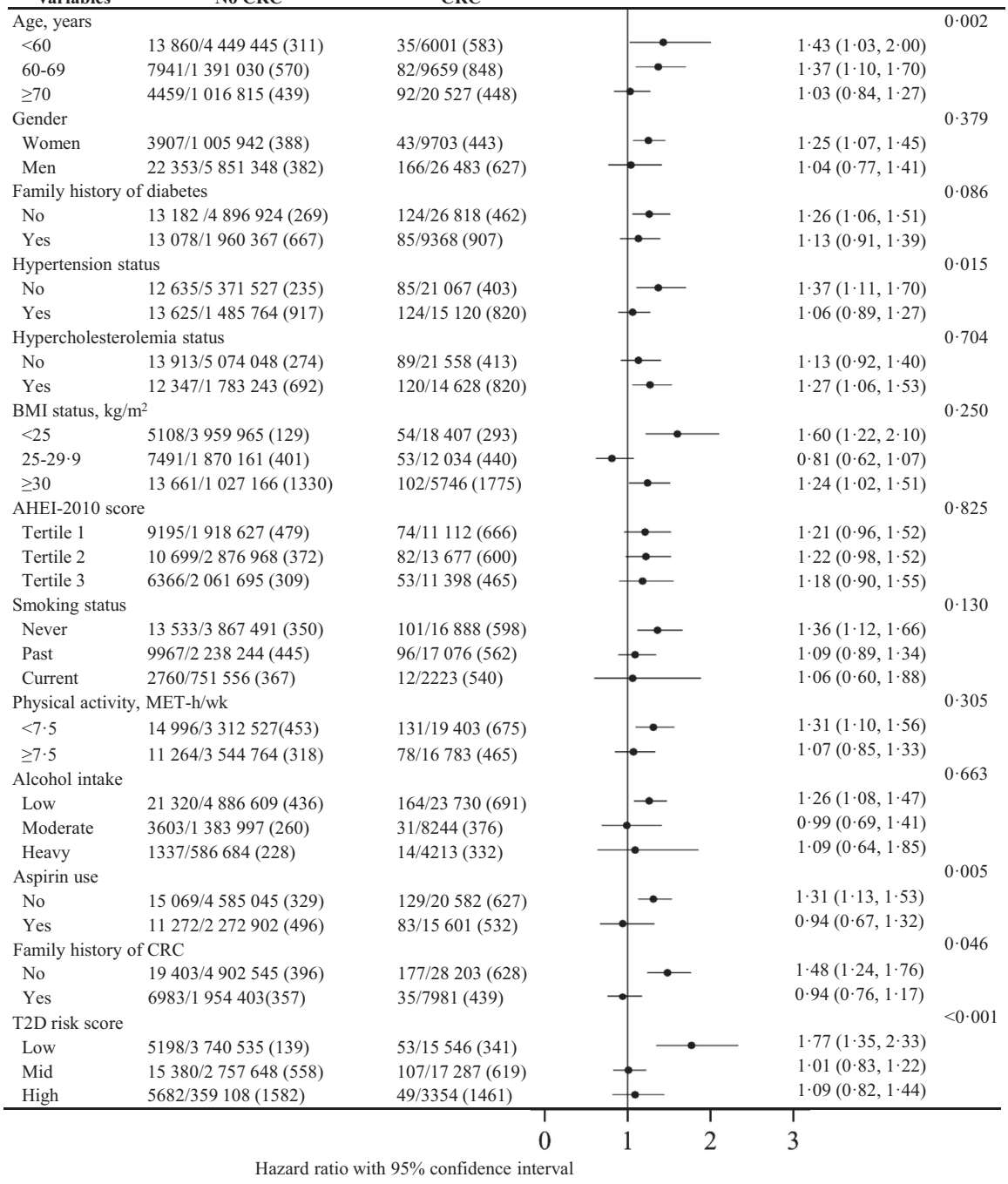

Fig. 2: HRs for T2D in subgroup analysis in the three cohorts.

Regardless of the underlying mechanism, our findings highlight the urgent need for improved recognition and diagnosis of T2D associated with CRC. It is important to establish practice guidelines for the prevention, diagnosis, and treatment of CRC-related T2D. Evidence shows that CRC survivors are less likely to receive appropriate care for diabetes compared to non-CRC cases.[39] We suspect the diagnosis and treatment of cancer may distract patients and health care providers from the appropriate prevention and management of

First author (year)	Country	Mean age (y)	Average follow-up years	No. of subjects	No. of CRC cases	No. of diabetes cases	Diabetes ascertainment	Covariate adjustment set	CRC duration (y)	Risk estimate HR (95% CI)
Khan (2011)	United Kingdom	74.1	10.1	25,196	5068	792	Medical records	BMI, comorbidity	Overall	1.39 (1.12, 1.72)
Singh (2016)	Canada	68	4.8	238,242	39707	20,765	Medical records	Neighborhood income quantile, rural residence, comorbidity index, hypertension, cardiovascular events, stroke, renal failure	Overall	1.31 (1.26, 1.36)
									≤1	1.53 (1.42, 1.64)
									1–2	1.36 (1.25, 1.47)
									2–3	1.31 (1.19, 1.43)
									3–4	1.19 (1.07, 1.33)
									4–5	1.19 (1.05, 1.35)
									5–6	1.15 (0.99, 1.34)
									6–7	1.13 (0.94, 1.35)
Hwangbo (2018)	South Korea	20–70[a]	7	494,189	1756	NA	Medical and diabetes drug records	Age, sex, BMI, smoking, alcohol intake, hypertension, hyperlipidemia, comorbidity index, systolic blood pressure, fasting glucose, total cholesterol	Overall	1.04 (0.86, 1.27)
Hawkins (2020)	The United States	63.7	8	33,093	7114	1239	Medical records	Birth year, sex, birth state, race, BMI, comorbidity index	Overall	1.14 (1.03,1.26)
									Diabetes with complication	1.19(0.98, 1.40)
									1–5	1.36 (1.09, 1.70)
									5.1–10	1.14 (0.84, 1.56)
									>10	0.84 (0.49, 1.47)
									Diabetes without complication	1.13 (0.99, 1.26)
									1–5	1.10 (0.94, 1.29)
									5.1–10	1.10 (0.88, 1.36)
									>10	1.39 (1.00, 1.93)
NHS	The United States	60.6	24.6	111,485	2079	13,833	Self-report and medical records verification	age, race, family history of heart disease, hypertension, smoking status, BMI, drinking status, multivitamin use, aspirin use, physical activity, family history of diabetes	Overall	1.19 (1.01, 1.44)
									≤5	1.10 (0.84, 1.44)
									5.1–10	1.23 (0.90, 1.67)
									>10	1.27 (0.95, 1.69)
NHS II	The United States	46.2	23.6	112,958	405	8686	Self-report and medical records	age, race, family history of heart disease, hypertension, smoking status, BMI, drinking status, multivitamin use, aspirin use, physical activity, family history of diabetes	Overall	1.74 (1.16, 2.60)
									≤5	1.45 (0.78, 2.69)
									5.1–10	2.03 (1.02, 4.06)
									>10	2.03 (0.91, 4.51)
HPFS	The United States	63.7	24.1	46,581	918	3950	Self-report and medical records verification	age, race, family history of heart disease, hypertension, smoking status, BMI, drinking status, multivitamin use, aspirin use, physical activity, family history of diabetes	Overall	1.04 (0.77, 1.41)
									≤5	1.18 (0.77, 1.79)
									5.1–10	1.06 (0.62, 1.83)
									>10	0.77 (0.39, 1.55)

[a]Age range.

Table 4: **Characteristics of studies included in the meta-analysis evaluating association of CRC and risk of T2D, 2011–2021.**

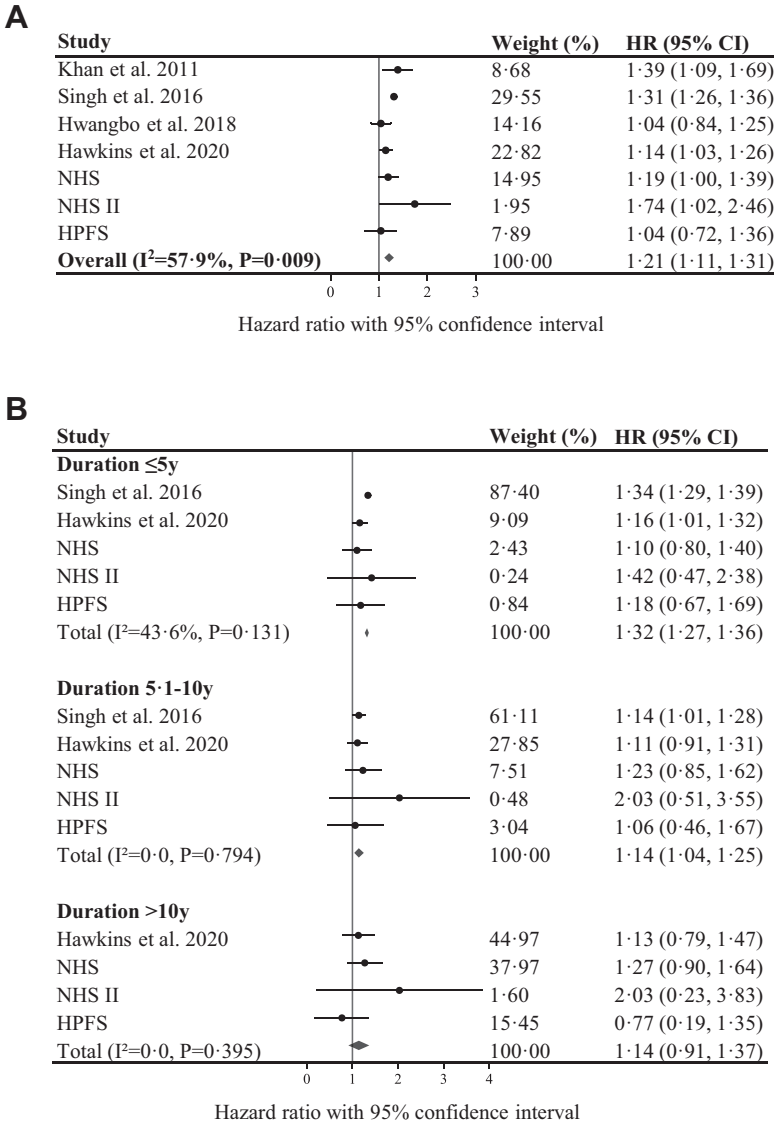

Fig. 3: Meta-analysis association between colorectal cancer (CRC) and risk of type 2 diabetes (T2D) overall and by CRC duration. A. Overall. B. By CRC duration.

other chronic diseases including T2D, or perhaps there are social determinants of health in terms of access to care at play.[40] Important to highlight is that many CRC survivors and their primary care providers are often unaware of this phenomenon, ultimately manifesting in negative impact of long-term health outcomes.[22,31] Improving awareness of this clinical scenario may help improve cancer patients' treatment safety and treatment efficacy. The stronger independent risk observed between CRC and T2D development in those without traditionally high T2D risk score implies that enhanced T2D prevention for CRC survivors is needed among all patients, regardless of presentation of traditional risk factors. Active surveillance of glucose, evaluation of metabolic status and assessing the impact of treatment on the development of metabolic system would fill a critical need for the prevention and recognition of the new onset of T2D.[41] Particularly, these prevention measures might be especially important in the first ten years after CRC diagnosis, based on our stratified results by CRC duration.

Some limitations of our study deserve further consideration. First, despite the overall large sample size with 6.9 million person-years follow-up, the number of T2D cases remained limited. Our study may be underpowered to test whether differences existed in some stratified analyses. Further research in larger populations is warranted to confirm conclusions and elucidate potential mechanisms.

Second, our cohorts' population primarily consisted of whites, possibly limiting the generalizability of findings to other ethnic groups, and to other countries with different population demographics. Furthermore, given the small number of T2D cases after CRC diagnosis in the NHS II and HPFS cohorts, the prospective results as well as meta-analysis results are primarily driven by the NHS I cohort. However, the consistent results from the systematic review and meta-analysis provide compelling evidence for a positive association between CRC and T2D in the US and globally. Based on the seven cohort studies with high study quality, the meta-analysis increased the statistical power and enabled us to rigorously assess the association between CRC and T2D, overall and by CRC duration. Moreover, the sensitivity analysis by exclusion of one study at a time yielded relatively stable risk estimates, indicating robustness of our findings. Third, we did not have complete treatment data in our study to examine the effect of treatment on subsequent T2D development. However, the persistent associations among subjects who had CRC for more than five years and early-stage CRC cases suggest that the association between CRC and T2D might, at least partly, be independent of cancer chemotherapy, which is consistent with previous reports.[6] Additional long-term prospective cohort studies with comprehensive information on individual-level treatment are needed to clarify the associations between treatment and late effect. Fourth, our study may be subject to detection bias if CRC patients are more likely to receive medical examination for symptoms. However, this seems unlikely, because when we excluded T2D cases within the first two years after CRC diagnosis, the results did not materially change. Meanwhile, the positive association was similarly observed in CRC patients during 5.1–10 years of diagnosis. Fifth, because the diagnostic information of T2D was collected at discrete time points, there may be measurement error in assessment of T2D duration.The T2D diagnostic criteria changed over timeour study may be subject to self-reporting bias. However, the questionnaires used in the cohorts have been comprehensively validated.[12,13,15,17,18,39–44] Finally, like other observational studies, we cannot rule out the possibility of unmeasured or unknown confounding.

In conclusion, we found that CRC diagnosis was associated with an increased risk of T2D, especially in the first ten years after CRC diagnosis. These findings highlight the need for larger studies and more in-depth research in this area, which will be helpful to advance biological understanding and improve prevention, early detection, and treatment of T2D after CRC diagnosis. Our study results indicate the importance of T2D prevention among CRC survivors.

Contributors

Each author's contribution was listed as follows. Hongmei Zeng: conceptualization, data curation, methodology, formal analysis, validation, funding, investigation, writing-original draft, writing-review & editing; Chen Yuan: conceptualization, methodology, validation, writing-original draft, writing-review & editing; Jakub Morze: conceptualization, data curation, methodology, formal analysis, validation, writing-original draft, writing-review & editing; Ruiying Fu: data curation, methodology, formal analysis, validation, writing-original draft, writing-review & editing; Kai Wang: methodology, writing-original draft, writing-review & editing; Liang Wang: methodology, validation, writing-original draft, writing-review & editing; Feng Sun: methodology, writing-original draft, writing-review & editing; John S Ji: methodology, writing-original draft, writing-review & editing. Edward L. Giovannucci: conceptualization, methodology, validation, writing-original draft, writing-review & editing, project administration; Mingyang Song: conceptualization, methodology, validation, writing-original draft, writing-review & editing, project administration. Hongmei Zeng, Chen Yuan, Jakub Morze, Ruiying Fu, Edward L. Giovannucci and Mingyang Song had directly accessed and verified the underlying data. Mingyang Song and Edward L. Giovannucci were responsible for the decision to submit the manuscript.

Data sharing statement

Due to the participant confidentiality and privacy concern, the cohorts' analytic data and data dictionary were not public available. Further information including the procedures to obtain and access data from the Nurses' Health Studies and Health Professionals Follow-up Study is described at https://www.nurseshealthstudy.org/researchers (contact email: nhsaccess@channing.harvard.edu) and https://sites.sph.harvard.edu/hpfs/for-collaborators/. All raw tabulated data from the systematic review and meta-analysis were available upon request to the corresponding authors.

Declaration of interests

We declare no competing interests.

Acknowledgements

We acknowledged the grants that were supporting the research: NHS cohort infrastructure grant (UM1 CA186107), NHS program project grant that funds cancer research (P01 CA87969), NHS II cohort infrastructure grant (U01 CA176726), HPFS cohort infrastructure grant (U01 CA167552), Cancer Hospital, Chinese Academy of Medical Sciences Hope Star Fund, and the Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2021-I2M-1-010).

We appreciate the participants and staff of the Nurses' Health Study, Nurses' Health Study II, and the Health Professionals Follow-up Study. We thank the staff at the following state cancer registries for their help. AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ebiom.2022.104345.

References

1 Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249.
2 McCabe MS, Bhatia S, Oeffinger KC, et al. American Society of Clinical Oncology statement: achieving high-quality cancer survivorship care. *J Clin Oncol*. 2013;31(5):631–640.
3 Khan NF, Mant D, Carpenter L, Forman D, Rose PW. Long-term health outcomes in a British cohort of breast, colorectal and prostate cancer survivors: a database study. *Br J Cancer*. 2011;105(Suppl 1):S29–S37.
4 Singh S, Earle CC, Bae SJ, et al. Incidence of diabetes in colorectal cancer survivors. *J Natl Cancer Inst*. 2016;108(6):djv402.
5 Hawkins ML, Blackburn BE, Rowe K, et al. Endocrine and metabolic diseases among colorectal cancer survivors in a population-based cohort. *J Natl Cancer Inst*. 2020;112(1):78–86.

6 Hwangbo Y, Kang D, Kang M, et al. Incidence of diabetes after cancer development: a Korean national cohort study. *JAMA Oncol*. 2018;4(8):1099–1105.
7 Giovannucci E, Harlan DM, Archer MC, et al. Diabetes and cancer: a consensus report. *Diabetes Care*. 2010;33(7):1674–1685.
8 Hu Y, Zhang X, Ma Y, et al. Incident type 2 diabetes duration and cancer risk: a prospective study in two US cohorts. *J Natl Cancer Inst*. 2021;113(4):381–389.
9 Belanger CF, Hennekens CH, Rosner B, Speizer FE. The nurses' health study. *Am J Nurs*. 1978;78(6):1039–1040.
10 Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann Surg Oncol*. 2010;17(6):1471–1474.
11 American Diabetes Association. Standards of medical care in diabetes–2010. *Diabetes Care*. 2010;33(Suppl 1):S11–S61.
12 Manson JE, Rimm EB, Stampfer MJ, et al. Physical activity and incidence of non-insulin-dependent diabetes mellitus in women. *Lancet*. 1991;338(8770):774–778.
13 Hu FB, Leitzmann MF, Stampfer MJ, Colditz GA, Willett WC, Rimm EB. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med*. 2001;161(12):1542–1548.
14 Stampfer MJ, Willett WC, Speizer FE, et al. Test of the national death index. *Am J Epidemiol*. 1984;119(5):837–839.
15 Chasan-Taber S, Rimm EB, Stampfer MJ, et al. Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology*. 1996;7(1):81–86.
16 Chiuve SE, Fung TT, Rimm EB, et al. Alternative dietary indices both strongly predict risk of chronic disease. *J Nutr*. 2012;142(6):1009–1018.
17 Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol*. 1994;23(5):991–999.
18 Colditz GA, Martin P, Stampfer MJ, et al. Validation of questionnaire information on risk factors and disease outcomes in a prospective cohort study of women. *Am J Epidemiol*. 1986;123(5):894–900.
19 Hosmer DW, Lemeshow S. *Applied survival analysis: regression modeling of time to event data*. New York: Wiley; 1999.
20 Griffin SJ, Little PS, Hales CN, Kinmonth AL, Wareham NJ. Diabetes risk score: towards earlier detection of type 2 diabetes in general practice. *Diabetes Metab Res Rev*. 2000;16(3):164–171.
21 Abbasi A, Peelen LM, Corpeleijn E, et al. Prediction models for risk of developing type 2 diabetes: systematic literature search and independent external validation study. *BMJ*. 2012;345:e5900.
22 Baliunas DO, Taylor BJ, Irving H, et al. Alcohol as a risk factor for type 2 diabetes: a systematic review and meta-analysis. *Diabetes Care*. 2009;32:2123–2132.
23 Dugee O, Janchiv O, Jousilahti P, et al. Adapting existing diabetes risk scores for an Asian population: a risk score for detecting undiagnosed diabetes in the Mongolian population. *BMC Public Health*. 2015;15:938.
24 Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed October 4, 2021.
25 Niederseer D, Bracher I, Stadlmayr A, et al. Association between cardiovascular risk and diabetes with colorectal neoplasia: a site-specific analysis. *J Clin Med*. 2018;7(12):1–20.
26 Wang M, Yang Y, Liao Z. Diabetes and cancer: epidemiological and biological links. *World J Diabetes*. 2020;11(6):227–238.
27 Oguntibeju OO. Type 2 diabetes mellitus, oxidative stress and inflammation: examining the links. *Int J Physiol Pathophysiol Pharmacol*. 2019;11(3):45–63.
28 Satoh K, Yachida S, Sugimoto M, et al. Global metabolic reprogramming of colorectal cancer occurs at adenoma stage and is induced by MYC. *Proc Natl Acad Sci U S A*. 2017;114(37):e7697–e7706.
29 Masi T, Patel BM. Altered glucose metabolism and insulin resistance in cancer-induced cachexia: a sweet poison. *Pharmacol Rep*. 2021;73(1):17–30.
30 Dev R, Bruera E, Dalal S. Insulin resistance and body composition in cancer patients. *Ann Oncol*. 2018;29(suppl_2):ii18–ii26.
31 Giovannucci E. The critical need for guidance in managing glycaemic control in patients with cancer. *Diabet Med*. 2021;39:e14624.
32 Tabak AG, Jokela M, Akbaraly TN, Brunner EJ, Kivimaki M, Witte DR. Trajectories of glycaemia, insulin sensitivity, and insulin secretion before diagnosis of type 2 diabetes: an analysis from the Whitehall II study. *Lancet*. 2009;373(9682):2215–2221.
33 Lee EK, Koo B, Hwangbo Y, et al. Incidence and disease course of new-onset diabetes mellitus in breast and colorectal cancer patients undergoing chemotherapy: a prospective multicenter cohort study. *Diabetes Res Clin Pract*. 2021;174:108751.
34 Ariaans G, de Jong S, Gietema JA, Lefrandt JD, de Vries EG, Jalving M. Cancer-drug induced insulin resistance: innocent bystander or unusual suspect. *Cancer Treat Rev*. 2015;41(4):376–384.
35 Bonaventura A, Montecucco F. Steroid-induced hyperglycemia: an underdiagnosed problem or clinical inertia? A narrative review. *Diabetes Res Clin Pract*. 2018;139:203–220.
36 Hwang JL, Weiss RE. Steroid-induced diabetes: a clinical and molecular approach to understanding and treatment. *Diabetes Metab Res Rev*. 2014;30(2):96–102.
37 Huang CY, Lin YS, Liu YH, Lin SC, Kang BH. Hyperglycemia crisis in head and neck cancer patients with platinum-based chemotherapy. *J Chin Med Assoc*. 2018;81(12):1060–1064.
38 Feng JP, Yuan XL, Li M, et al. Secondary diabetes associated with 5-fluorouracil-based chemotherapy regimens in non-diabetic patients with colorectal cancer: results from a single-centre cohort study. *Colorectal Dis*. 2013;15(1):27–33.
39 Griffiths RI, Keating NL, Bankhead CR. Quality of diabetes care in cancer: a systematic review. *Int J Qual Health Care*. 2019;31(2):75–88.
40 Arreskov AB, Olsen MA, Pouplier SS, et al. The impact of cancer on diabetes outcomes. *BMC Endocr Disord*. 2019;19(1):60.
41 Siegel KR, Ali MK, Zhou X, et al. Cost-effectiveness of interventions to manage diabetes: has the evidence changed since 2008? *Diabetes Care*. 2020;43(7):1557–1592.
42 Al-Shaar L, Yuan C, Rosner B, et al. Reproducibility and validity of a semiquantitative food frequency questionnaire in men assessed by multiple methods. *Am J Epidemiol*. 2021;190(6):1122–1132.
43 Troy LM, Hunter DJ, Manson JE, Colditz GA, Stampfer MJ, Willett WC. The validity of recalled weight among younger women. *Int J Obes Relat Metab Disord*. 1995;19(8):570–572.
44 Yuan C, Spiegelman D, Rimm EB, et al. Validity of a dietary questionnaire assessed by comparison with multiple weighed dietary records or 24-hour recalls. *Am J Epidemiol*. 2017;185(7):570–584.