IJC
International Journal of Cancer

Short Report

Diabetes mellitus, genetic variants in the insulin-like growth factor pathway and colorectal cancer risk

Sander de Kort[1,2,3], Colinda C.J.M. Simons[1], Piet A. van den Brandt[1], Maryska L.G. Janssen-Heijnen[1,4], Silvia Sanduleanu[2,3], Ad A.M. Masclee[2,3] and Matty P. Weijenberg[1]

[1]Department of Epidemiology, GROW—School for Oncology and Developmental Biology, Maastricht University, Maastricht, The Netherlands
[2]Division of Gastroenterology and Hepatology, Department of Internal Medicine, Maastricht University Medical Centre, Maastricht, The Netherlands
[3]GROW—School for Oncology and Developmental Biology, NUTRIM, School for Nutrition, Toxicology and Metabolism, Maastricht University, Maastricht, The Netherlands
[4]Department of Clinical Epidemiology, VieCuri Medical Centre, Venlo, The Netherlands

Genetic variation in the insulin-like growth factor (IGF) pathway may further increase the risk of colorectal cancer (CRC) associated with type 2 diabetes mellitus (T2DM). Joint effects of T2DM and genetic variation in the IGF pathway on CRC risk can increase mechanistic insights. Participants from the Netherlands Cohort Study (n = 120, 852) completed a baseline questionnaire in 1986 when 55–69 years old (case–cohort, $n_{subcohort}$ = 5,000, n_{cases} = 3,441 after 16.3 years follow-up). Self-reported DM at baseline with onset at ≥30 years was classified as T2DM. Eighteen single nucleotide polymorphisms (SNPs) from the IGF pathway were aggregated in a genetic risk score (GRS). Cox proportional hazard ratios (HRs) for CRC were estimated according to combinations of T2DM status with GRS tertiles and categories of an *IGF1* 19-CA repeat polymorphism. Baseline T2DM prevalence was 3.1% in subcohort members and 3.8% in CRC cases. Comparison of combined categories with non-T2DM individuals in the lowest GRS tertile as reference showed that those in the highest GRS tertiles with and without T2DM had significantly increased CRC risks, particularly those with T2DM (HR = 2.28, 95% CI: 1.11, 4.66). As compared to *IGF1* 19-CA wild-type carriers without T2DM, carrying two *IGF1* 19-CA variant repeat alleles were associated with a significantly decreased CRC risk in those without T2DM (HR = 0.76, 95% CI: 0.63–0.91). This association was absent when T2DM was present. Our study of joint effects indicated that the presence of unfavorable alleles in the IGF pathway may further increase the risk of CRC associated with T2DM.

Introduction

Colorectal cancer (CRC) and type 2 diabetes mellitus (T2DM) are multifactorial diseases that have several risk factors in common (e.g., high BMI, physical inactivity) and both have shown an increase in incident rates over the past decades.[1,2] T2DM has been associated with an increased CRC risk (20–40%).[3] An interesting hypothesis underlying this association involves the insulin-like growth factor (IGF) pathway. In this hypothesis, chronic hyperinsulinemia in T2DM patients leads to increased IGF levels, which can accelerate the progression from adenoma to cancer.[4] Previous research within the Netherlands Cohort Study (NLCS) studied single nucleotide polymorphisms (SNPs) in the IGF pathway in relation to CRC risk and found that carrying more risk alleles was associated with an increased CRC risk in men, specifically an increased proximal and distal colon cancer risk but not rectal cancer risk.[5] T2DM was associated with an increased proximal colon cancer risk in women in previously published data within the NLCS[6] and this confirmed findings

Key words: cohort studies, colorectal cancer, insulin-like growth factors, gene–environment interactions, diabetes mellitus

Abbreviations: BMI: body mass index; CRC: colorectal cancer; CIs: confidence intervals; GRS: genetic risk score; HWE: Hardy–Weinberg equilibrium; HRs: hazard ratios; IGF: insulin-like growth factor; NLCS: Netherlands Cohort Study; RERI: relative excess risk due to interaction; SNPs: single nucleotide polymorphisms; T2DM: type 2 diabetes mellitus

Conflict of interest: The authors have no conflict of interest to declare.

Grant sponsor: Health Foundation Limburg; **Grant sponsor:** Dutch Cancer Society; **Grant number:** 2009-4281

DOI: 10.1002/ijc.32365

History: Received 18 Oct 2018; Accepted 2 Apr 2019; Online 24 Apr 2019

Correspondence to: Sander de Kort Division of Gastroenterology and Hepatology, Department of Internal Medicine Maastricht University Medical Centre, PO Box 5800, 6202 AZ, Maastricht, The Netherlands, Tel.: +31-433-875-021, Fax: +31 433 875 006, E-mail: s.dekort@maastrichtuniversity.nl

What's new?

Colorectal cancer (CRC) and type-2 diabetes mellitus (T2DM) have several risk factors in common (high BMI, physical inactivity, etc.). In this study, the authors identified several SNPs in the insulin-like growth factor (IGF) pathway that further increase CRC risk, particularly in people with T2DM. These results support the hypothesis that T2DM increases CRC risk by influencing the IGF pathway. Screening for these unfavorable genetic variations in the IGF pathway may also help to improve the assessment of personalized CRC risk.

from other cohort studies.[7] Here, we aimed to investigate whether carrying unfavorable alleles in the IGF pathway further increased the CRC risk associated with diabetes. Genetic variants were used as markers of pathway involvement. Finding a further increased CRC risk in diabetics with unfavorable alleles in the IGF pathway would substantiate the hypothesis that diabetes or chronic hyperinsulinemia increases CRC risk by influencing the IGF pathway.

Materials and Methods

Study population and design

The NLCS includes 120,852 men and women from the Dutch population, who were 55–69 years old at baseline in 1986 when completing a self-administered questionnaire on diet and cancer.[8] Roughly 90,000 participants also provided toenail clippings when returning the baseline questionnaire. Toenail DNA is a valid DNA source for genotyping.[9] For reasons of efficiency, DNA isolation, the processing of questionnaires and follow-up, the NLCS is characterized by a case–cohort approach. This approach entails that a random subcohort of 5,000 individuals, selected at random immediately after baseline, is followed-up for vital status to estimate the accumulated person-time at risk by record linkage to the municipal registries (>99% complete). The whole cohort is followed-up for incident cancer by record linkage to the Dutch cancer registry and the national pathology database (PALGA)[10] (>96% complete).[11] Baseline exclusion of participants with a history of cancer (except skin cancer) left 4,774 subcohort members. After 16.3 years, there were 3,441 incident CRC cases (ICD-O 153–154). Toenail clippings were available for 3,768 of 4,774 subcohort members (78.9%) and 2,580 of 3,441 CRC cases (75.0%). Finally, after exclusion of missing inconsistent / incomplete questionnaires and missing data on diabetes and covariates, 1907 CRC cases, 565 proximal colon cancer cases (ICD-O 153.0, 153.1, 153.4, 153.5 and 153.6), 533 distal colon cancer cases (codes 153.2, 153.3 and 153.7) and 432 rectal cancer cases (code 154.1) were available for analyses with respect to the genetic risk score (GRS).

Assessment of T2DM status

T2DM status was derived from the baseline questionnaire based on the question: "Has a physician ever diagnosed you with diabetes mellitus and what was your age at that time?" Individuals who reported to have been diagnosed with DM at age 30 years or older were classified as T2DM. This cut-off point was based on epidemiological data on DM prevalence and previous literature.[12] Individuals with a diagnosis of DM before the age of 30 years were regarded as non-T2DM in our study.

Variant selection, genotyping, the GRS and the *IGF1* 19-CA repeat

We selected SNPs in genes encoding for factors in or regulatory to the IGF pathway and genes encoding for adiponectin, adiponectin receptors and peroxisome proliferator-activated receptor gamma, as these hormones regulate insulin resistance. Briefly, SNPs were selected through literature and had to be associated with CRC or other obesity-related cancers (e.g., breast cancer), T2DM risk or related traits (e.g., body size or markers of insulin resistance). A full description of our gene and SNP selection strategy, including a list of the selected SNP variants and references to the literature on which selection is based, is available in the paper and Supplemental Material of Simons *et al.*[5] A total of 24 SNPs were successfully genotyped on the SEQUENOM® MassARRAY® system (Hamburg, Germany; now known as Agena Bioscience, Hamburg, Germany) and an *IGF1* 19-CA repeat polymorphism was genotyped by PCR amplification and subsequent analysis of the PCR products' length using the 96-capillary ABI 3730xl DNA Analyzer. SNP genotypes were reproducible in 98.8% of 314 random duplicate samples. SNP call rates were 92.6% or higher, with one exception of 83.6% (rs4773082). Twenty-three out of 24 SNPs adhered to Hardy–Weinberg equilibrium (HWE). We did not exclude the SNP that deviated from HWE (rs1342387) because one in 20 tests may be expected to be significant by chance alone. We only used samples with a sample call rate of >95% in statistical analyses, which led to the exclusion of 532 samples. We aggregated 18 SNPs into a GRS, as for these SNPs, the risk allele in relation to the endpoints mentioned above was unequivocal on the basis of previous studies. Which allele (minor or major allele) was the risk allele can be found in the Supplemental Material of Simons *et al.*[5] Per SNP, individuals carry 2 (homozygotes for the risk allele), 1 (heterozygotes) or 0 (homozygotes for the other allele) risk alleles. The number of risk alleles carried by an individual was summed for the 18 SNPs. The theoretical range of a GRS based on 18 SNPs is between 0 and 36 risk alleles. The GRS was categorized into tertiles based on the distribution in the subcohort. The *IGF1* 19-CA repeat polymorphism had a call rate of 70.7% and was reproducible in 93.6% of 314 duplicate samples. This variant was analyzed separately from the GRS, as the direction of the association with CRC remains uncertain and because it is a conceptually different variant (not a single base pair change, but a variation in

the number of CA repeats). According to Rosen *et al.*,[13] we distinguished between individuals homozygous for the wild-type allele (192/192 CA repeats), heterozygous individuals (192/non-192 CA repeats), and individuals carrying two variant alleles (non-192/non-192 CA repeats).

Statistical analysis

Hazard ratios (HRs) and 95% confidence intervals (CIs) for CRC overall, by sex (only CRC), and by subsite were estimated using Cox proportional hazards regression analysis for combined categories of T2DM status with tertiles of the GRS and categories of the *IGF1* 19-CA repeat polymorphism. To adjust for the additional variance introduced from sampling the subcohort from the total cohort, standard errors were estimated using the robust Huber–White sandwich estimator.[14] Exclusion of participants with inconsistent/incomplete baseline questionnaires left 2,729 subcohort members and 1,821 CRC cases which could be categorized in one of the tertiles of the GRS, and 2,042 subcohort members and 1,657 CRC cases with *IGF1* 19-CA repeat status available for analyses. Models were adjusted for potential confounders [age (years), body mass index (BMI, kg/m^2), pant/skirt size as a proxy for abdominal circumference[15] (below, or equal to or above median size), first-degree family history of CRC (yes/no), smoking status (never, ex, current), alcohol intake (0, 0.1–29, ≥30 g/day), intake in g/day of total meat, vegetables, fish, sweets, added sugar, saturated fats and fiber, total energy intake (kcal/day), and nonoccupational physical activity (≤30, >30, >60, >90 min/day)]. The relative excess risk due to interaction (RERI) was calculated according to Rothman *et al.*,[16] though we conservatively refrained from doing so in analyses including the *IGF1* 19-CA repeat polymorphism, as the combined category at lowest risk was not clear. Corresponding 95% bias-corrected confidence intervals for the RERI were estimated by bootstrapping (n bootstrap samples = 1,000).[17] Multiplicative interactions were assessed using the Wald test. All analyses were conducted using Stata version 12 (Stata Corp., College Station, TX). Statistical significance was indicated by a p-value <0.05 for two-sided testing.

Results

Table 1 shows the distribution of subcohort members with and without T2DM across categories of age, sex, the GRS, the *IGF1* 19-CA repeat polymorphism and potential confounders. Compared to subcohort members without T2DM, subcohort members with T2DM were older ($p < 0.001$) and had a significantly lower intake of sweets and added sugar ($p < 0.001$ for both).

The GRS and T2DM

CRC risks were increased across combined categories of T2DM and GRS using non-T2DM individuals in the lowest tertile of the GRS as reference (Table 2). Individuals with T2DM in the lowest GRS tertile did not differ in CRC risk (HR = 1.00, 95% CI: 0.54–1.83) from those without T2DM in the lowest tertile. Non-T2DM individuals in the middle and highest GRS tertiles were at a statistically significant increased CRC risk as compared to non-T2DM individuals in the lowest GRS tertile ($HR_{\text{non-T2DM in middle GRS tertile } vs. \text{ non-T2DM in lowest GRS tertile}}$ = 1.23, 95% CI: 1.06–1.42 and $HR_{\text{non-T2DM in highest GRS tertile } vs. \text{ non-T2DM in lowest GRS tertile}}$ = 1.27, 95% CI: 1.09–1.49), as were T2DM individuals in the middle and highest GRS tertiles $HR_{\text{T2DM in middle GRS tertile } vs. \text{ non-T2DM in lowest GRS tertile}}$ = 1.82, 95% CI: 1.09–3.05, and $HR_{\text{T2DM in highest GRS tertile } vs. \text{ non-T2DM in lowest GRS tertile}}$ = 2.28, 95% CI: 1.11–4.66). Sex-specific analyses showed similar results as the overall analysis, although most HRs were not statistically significant. Subsite-specific analyses showed that the results seen for CRC overall were most pronounced for proximal colon cancer ($HR_{\text{T2DM in highest GRS tertile } vs. \text{ non-T2DM in lowest GRS tertile}}$ = 3.92, 95% CI: 1.75–8.79), but less clear in relation to distal colon cancer and rectal cancer. Power was insufficient to investigate subsite-specific CRC risks in men and women separately. Stratified analyses yielded results that were not statistically significant, except when considering proximal colon cancer as endpoint and comparing T2DM with non-T2DM individuals within the highest GRS tertile (HR = 3.93, 95% CI: 1.63–9.48). Tests for multiplicative and additive interactions (RERI) were not statistically significant.

The *IGF1* 19-CA repeat polymorphism and T2DM

Table 3 shows the CRC risks associated with combined categories of T2DM and the *IGF1* 19-CA repeat polymorphism. Using non-T2DM individuals homozygous for the wild-type 19-CA repeat allele as reference showed that non-T2DM individuals carrying one or two variant *IGF1* 19-CA repeat alleles had (a borderline) statistically significantly decreased CRC risk (HR = 0.88, 95% CI: 0.75–1.04 and HR = 0.76, 95% CI: 0.63–0.91, respectively), whereas CRC risk was not significantly changed in individuals with T2DM homozygous for the wild-type *IGF1* 19-CA repeat allele or those carrying one or two variant alleles (HR = 1.44, 95% CI: 0.76–2.74; HR = 0.73, 95% CI: 0.37–1.44; and HR = 1.26, 95% CI: 0.53–2.96, respectively). Sex-specific analyses showed similar results in women and nonsignificant results in men. Subsite-specific analyses showed similar results as overall analyses. Stratified analyses showed no statistically significant associations between T2DM and CRC risk overall or by subsite, except when considering proximal colon cancer as endpoint and comparing T2DM with non-T2DM individuals within those carrying two variant IGF1 19-CA repeat alleles (HR = 3.09, 95% CI: 1.03–9.24). Multiplicative and additive interactions were not statistically significant.

Discussion

In this long-term prospective study, we used genetic variation in the IGF pathway as time-independent markers of IGF pathway involvement. To the best of our knowledge, our study is the first to simultaneously consider T2DM with markers of IGF pathway involvement in relation to CRC risk. Comparison of combined categories with non-T2DM in the lowest GRS tertile as reference showed that the presence of more unfavorable alleles in the IGF

Table 1. Baseline characteristics of subcohort members and CRC cases with and without type 2 diabetes mellitus in the Netherlands Cohort Study

	Subcohort members			CRC Cases		
Characteristic	**No T2DM (*n* = 2,648; 97%)**	**T2DM (*n* = 81; 3.1%)**	***p***	**No T2DM (*n* = 1,754; 96%)**	**T2DM (*n* = 67; 3.8%)**	***p***
Age at baseline, mean in years (SD)	61.3 (4.2)	63.5 (4.3)	‹0.001	62.0 (4.1)	63.0 (3.8)	0.03
Age at CRC diagnosis, mean in years (SD)				71.2 (5.6)	71.3 (5.6)	0.81
Gender (% male)	1,318 (50)	41 (51)	0.88	1,047 (60)	32 (48)	0.05
Genetic sum score, *n* (%)						
Tertile 1	988 (37)	34 (42)		557 (32)	18 (27)	
Tertile 2	991 (38)	34 (42)		697 (40)	33 (49)	
Tertile 3	669 (25)	13 (16)	0.17	500 (28)	16 (24)	0.30
IGF1 CA repeat[1], *n* (%)						
Non-19/non-19	509 (19)	11 (14)		341 (19)	11 (17)	
19/non-19	684 (26)	24 (30)		566 (32)	14 (21)	
19/19	600 (23)	18 (22)	0.60	522 (30)	25 (37)	0.12
BMI, mean kg/m^2 ± SD	25.0 (3.1)	25.6 (3.6)	0.14	25.0 (2.9)	25.6 (3.2)	0.09
Pants/skirt size, *n* (%) equal or above median	1,568 (59)	53 (65)	0.26	1,086 (62)	50 (75)	0.04
Family history of CRC, *n* (%) yes	160 (6)	3 (4)	0.63	176 (10)	4 (6)	0.27
Smoking Status, *n* (%)						
Never	917 (35)	35 (43)		525 (30)	29 (43)	
Ex-smoker	1,002 (38)	30 (37)		769 (44)	28 (42)	
Current smoker	729 (27)	16 (20)	0.18	460 (26)	10 (15)	0.03
Alcohol intake, *n* (%)						
0 g/day	610 (23)	23 (29)		342 (20)	25 (37)	
0.1–30 g/day	1,801 (68)	48 (59)		1,181 (67)	37 (55)	
›30 g/day	237 (9)	10 (12)	0.24	231 (13)	5 (8)	0.001
Dietary habits						
Total meat, mean in g/day (SD)[2]	113 (46)	120 (46)	0.40	115 (47)	122 (41)	0.10
Vegetables, mean in g/day (SD)	381 (159)	384 (148)	0.86	376 (164)	412 (153)	0.03
Fish, mean in g/day (SD)	13 (15)	13 (17)	0.81	12 (15)	13 (13)	0.48
Sweets, mean in g/day (SD)[3]	42 (26)	31 (27)	‹0.001	42 (26)	31 (25)	‹0.001
Added sugar, mean in g/day (SD)	20 (27)	5 (17)	‹0.001	21 (28)	1 (4)	‹0.001
Saturated fats, mean in g/day (SD)	33 (11)	31 (11)	0.13	34 (11)	36 (16)	0.17
Fiber intake, mean in g/day (SD)	27 (8)	27 (8)	0.51	27 (8)	29 (9)	0.10
Energy, mean in kcal/day (SD)	1,922 (505)	1,823 (456)	0.06	1,976 (495)	1,928 (588)	0.51
Physical activity, *n* (%)						
≤30 min/day	512 (19)	21 (26)		340 (19)	19 (20)	
›30–≤60 min/day	860 (33)	26 (32)		533 (31)	16 (24)	
›60–≤90 min/day	566 (21)	12 (15)		386 (22)	18 (27)	
›90 min/day	710 (27)	22 (27)	0.34	495 (28)	14 (21)	0.14

[1]Numbers of IGF1 19-CA repeat in baseline and analyses do not match due to exclusion of missing on the genetic sum score variable.
[2]Consists of the sum of pork, beef, poultry, game, and processed meat intakes.
[3]Consists of the sum of intakes of sweet sandwich filling, cookies, cake, and candy.
Abbreviations: BMI, body mass index; CRC, colorectal cancer; SD standard deviation; T2DM, type 2 diabetes mellitus.

pathway was associated with an increased CRC risk in both the presence and absence of T2DM, with strongly increased CRC risks observed in the presence of T2DM. These data suggest the main effect of the GRS and a possible joint effect of T2DM and the GRS. As compared to non-T2DM wild-type *IGF1* 19-CA repeat carriers, carrying two *IGF1* 19-CA variant repeat alleles was associated with a decreased CRC risk in the absence of T2DM, and this association was absent in the presence of T2DM. However, in all analyses, multiplicative and additive interactions were not statistically significant. We found no clear sex

Table 2. Multivariable-adjusted hazard ratios for colorectal cancer overall and by sex and subsite for combined categories of type 2 diabetes mellitus and tertiles of genetic risk score based on SNPs in genes from the IGF pathway and for type 2 diabetics *vs.* nondiabetics within tertiles of this genetic risk score in the Netherlands Cohort Study after 16.3 years of follow-up

		Type 2 diabetes mellitus										Type 2 diabetes mellitus		
		No				Yes						Yes *vs.* No		*p*-Value for multiplicative interaction
		PY	N cases	HR[1]	(95% CI)	PY	N cases	HR[1]	(95% CI)	RERI[1]	(95% CI)	HR[1]	(95% CI)	
Colorectum	T1	14,242	587	1	(reference)	415	19	1.00	(0.54–1.83)			1.10	(0.59–2.07)	
Genetic risk score[2]	T2	14,176	732	1.23	(1.06–1.42)	446	34	1.82	(1.09–3.05)			1.44	(0.85–2.42)	
	T3	9,736	517	1.27	(1.09–1.49)	183	18	2.28	(1.11–4.66)	1.01	(–0.46–4.28)	1.98	(0.93–4.19)	0.43
Colorectum men	T1	6,821	346	1	(reference)	238	8	0.60	(0.25–1.46)			0.65	(0.26–1.62)	
Genetic risk score	T2	6,951	431	1.21	(0.99–1.47)	194	18	1.94	(0.92–4.09)			1.50	(0.69–3.27)	
	T3	4,428	316	1.41	(1.13–1.75)	66	9	2.06	(0.71–5.95)	1.05	(–1.08–4.97)	1.43	(0.47–4.36)	0.22
Colorectum women	T1	7,421	241	1	(reference)	177	11	1.79	(0.77–4.17)			2.05	(0.86–4.92)	
Genetic risk score	T2	7,225	301	1.25	(1.00–1.55)	252	16	1.96	(0.94–4.07)			1.61	(0.76–3.41)	
	T3	5,308	201	1.15	(0.91–1.47)	118	9	2.14	(0.78–5.89)	0.19	(–3.04–5.80)	1.96	(0.66–5.77)	0.96
Proximal colon cancer	T1	14,242	207	1	(reference)	415	8	1.17	(0.52–2.65)			1.35	(0.57–3.21)	
Genetic risk score	T2	14,176	243	1.15	(0.93–1.42)	446	14	2.09	(1.07–4.08)			1.74	(0.88–3.43)	
	T3	9,736	163	1.14	(0.90–1.44)	183	12	3.92	(1.75–8.79)	2.61	(–0.47–7.17)	3.93	(1.63–9.48)	0.18
Distal colon cancer	T1	14,242	190	1	(reference)	415	5	0.79	(0.30–2.13)			0.78	(0.28–2.15)	
Genetic risk score	T2	14,176	236	1.22	(0.98–1.52)	446	8	1.42	(0.63–3.17)			1.10	(0.48–2.50)	
	T3	9,736	177	1.36	(1.08–1.72)	183	4	1.53	(0.48–4.90)	0.37	(–1.66–3.84)	1.13	(0.33–3.91)	0.83
Rectal cancer	T1	14,242	132	1	(reference)	415	4	0.88	(0.29–2.62)			0.95	(0.30–2.98)	
Genetic risk score	T2	14,176	167	1.24	(0.97–1.60)	446	7	1.72	(0.73–4.08)			1.40	(0.58–3.35)	
	T3	9,736	120	1.36	(1.03–1.78)	183	2	1.25	(0.27–5.92)	0.02	(–1.70–6.02)	0.92	(0.18–4.79)	0.78

[1]Hazard ratios were adjusted for age (years), body mass index (kg/m^2), pants/skirt size (below, or equal to or above median size), family history of colorectal cancer (yes/no), smoking status (never, ex, current), alcohol intake (0, 0.1–29, ≥30 g/day), intake in g/day of total meat, vegetables, fish, sweets, added sugar, saturated fats, and fiber, total energy intake (kcal/day) and nonoccupational physical activity (≤30, ›30, ›60 and ›90 min/day).

[2]The range in tertiles of the (literature-based) genetic risk score was 6–14, 15–18 and 19–29 risk alleles. The theoretical maximum was 36. The distribution of subcohort members across tertiles was 988 (37.3%), 991 (37.4%) and 669 (25.3%), respectively, in individuals without type 2 diabetes mellitus and 34 (42.0%), 34 (42.0%) and 13 (16.0%), respectively, in individuals with type 2 diabetes mellitus.

Abbreviations: CI, confidence interval; HR, hazard ratio; PY, person-years at risk; RERI, relative excess risk due to interaction.

Table 3. Multivariable-adjusted hazard ratios for colorectal cancer overall and by sex and subsite for combined categories of type 2 diabetes mellitus and the *IGF1* 19-CA repeat polymorphism and for type 2 diabetics *vs.* non-diabetics within categories of the *IGF1* 19-CA repeat polymorphism in the Netherlands Cohort Study after 16.3 years of follow-up

		Type 2 diabetes mellitus										Type 2 diabetes mellitus		
		No				Yes						Yes *vs.* No		*p*-Value for multiplicative interaction
		PY	N cases	HR[1]	(95% CI)	PY	N cases	HR[1]	(95% CI)	RERI[1]	(95% CI)	HR[1]	(95% CI)	
Colorectum	192/192	9,296	620	1	(reference)	280	26	1.44	(0.76–2.74)			1.44	(0.76–2.72)	
IGF1 19-CA repeat[2]	192/non-192	10,950	659	0.88	(0.75–1.04)	304	16	0.73	(0.37–1.44)			0.83	(0.42–1.65)	
	Non-192/non-192	8,238	389	0.76	(0.63–0.91)	153	12	1.26	(0.53–2.96)	0.05	(−1.73–1.84)	1.89	(0.75–4.75)	0.36
Colorectum men	192/192	5,027	371	1	(reference)	115	10	1.05	(0.38–2.89)			1.01	(0.34–2.98)	
IGF1 19-CA repeat	192/non-192	5,628	389	0.90	(0.73–1.12)	184	8	0.57	(0.23–1.43)			0.57	(0.23–1.45)	
	Non-192/non-192	3,274	234	0.98	(0.76–1.25)	48	5	1.57	(0.37–6.62)	0.54	(−2.57–18.4)	1.88	(0.39–8.94)	0.52
Colorectum women	192/192	4,269	249	1	(reference)	165	16	1.87	(0.83–4.18)			1.96	(0.84–4.54)	
IGF1 19-CA repeat	192/non-192	5,322	270	0.85	0.83–1.08	120	8	1.22	(0.44–3.42)			1.07	(0.34–3.34)	
	Non-192/non-192	4,964	155	0.54	0.42–0.71	105	7	1.21	(0.41–3.60)	−0.20	(−3.85–1.75)	2.35	(0.79–6.97)	0.85
Proximal colon cancer	192/192	9,296	218	1	(reference)	280	13	1.95	(0.91–4.16)			1.97	(0.90–4.30)	
IGF1 19-CA repeat	192/non-192	10,950	229	0.86	(0.69–1.07)	304	9	1.21	(0.53–2.74)			1.45	(0.63–3.34)	
	Non-192/non-192	8,238	123	0.64	(0.50–0.83)	153	6	1.63	(0.58–4.60)	0.04	(−3.00–3.07)	3.09	(1.03–9.24)	0.66
Distal colon cancer	192/192	9,296	215	1	(reference)	280	5	0.81	(0.29–2.27)			0.77	(0.27–2.16)	
IGF1 19-CA repeat	192/non-192	10,950	209	0.82	(0.65–1.02)	304	3	0.41	(0.12–1.39)			0.51	(0.15–1.74)	
	Non-192/non-192	8,238	131	0.76	(0.58–0.96)	153	4	1.32	(0.40–4.36)	0.76	(−1.27–2.79)	1.95	(0.54–7.07)	0.34
Rectal cancer	192/192	9,296	126	1	(reference)	280	5	1.44	(0.51–4.09)			1.53	(0.54–4.37)	
IGF1 19-CA repeat[2]	192/non-192	10,950	147	0.97	(0.74–1.27)	304	3	0.64	(0.18–2.25)			0.65	(0.18–2.35)	
	Non-192/non-192	8,238	91	0.90	(0.66–1.22)	153	2	1.07	(0.23–5.04)	−0.28	(−2.65–2.10)	1.30	(0.27–6.35)	0.64

[1]Hazard ratios were adjusted for age (years), body mass index (kg/m^2), pants/skirt size (median split, Dutch sizes), family history of colorectal cancer (yes/no), smoking status (never, ex, current), alcohol intake (0, 0.1–29, ≥30 g/day), intake in g/day of total meat, vegetables, fish, sweets, added sugar, saturated fats and fiber, total energy intake (kcal/day) and nonoccupational physical activity (≤30, >30, >60 and >90 min/day).

[2]The distribution of subcohort members across categories of the *IGF1* 19-CA repeat polymorphism was 567 (28.6%), 751 (37.8%) and 667 (33.6%), respectively, in individuals without type 2 diabetes mellitus and 13 (35.0%), 24 (42.0%) and 20 (23.0%), respectively, in individuals with type 2 diabetes mellitus.

Abbreviations: CI, confidence interval; HR, hazard ratio; RERI, relative excess risk due to interaction; PY, person-years; IGF1, insulin-like growth factor 1.

Int. J. Cancer: **145**, 1774–1781 (2019) © 2019 The Authors. *International Journal of Cancer* published by John Wiley & Sons Ltd on behalf of UICC

differences in any of the analyses performed. Subsite-specific analyses in relation to proximal colon cancer risk were most similar to overall analyses.

The most recent meta-analysis has shown T2DM to be a risk factor for CRC,[3] although the individual studies used, showed conflicting results regarding subsite- and sex-specific CRC risks.[18,19] Using NLCS data,[6] we have previously reported T2DM to be associated with an increased proximal colon cancer risk in women. In the current study, findings should be interpreted with caution as the number of CRC cases with T2DM per CRC subsite was low. However, the finding that T2DM was a statistically significant risk factor for proximal colon cancer in the presence of an accumulation of risk alleles in the IGF pathway at least shows consistency with previous findings regarding the T2DM-CRC association in the NLCS and other cohort studies.[7] Differences in embryological origin of subsites and consequently variations in susceptibility to influences of T2DM and IGF1 as a growth factor could theoretically explain why we observed associations with proximal colon cancer specifically.[20]

The absence of statistically significant interactions between T2DM and a GRS based on risk alleles in genes in the IGF pathway may be due to a lack of power resulting from the low prevalence of T2DM in our study and the absence of a (strong) main effect of T2DM on CRC risk overall in individuals in the lowest GRS tertile. A prior study by Simons *et al.*[21] within our research group showed that larger body size may be a CRC risk factor in men in the presence of an accumulation of unfavorable alleles in the IGF pathway. This previous study gives confidence in our present findings as both BMI and T2DM are independent CRC risk factors that both are thought to act through the IGF pathway.[22]

The analyses in which we simultaneously considered T2DM and the *IGF1* 19-CA repeat polymorphism showed no risk pattern other than that *IGF1* 19-CA variant repeat carriers as compared to wild-type repeat carriers in the absence of T2DM were associated with a (borderline) statistically significantly lower CRC risk. Although there is some inconsistency in the literature regarding the direction of the association between the *IGF1* 19-CA repeat polymorphism and CRC risk,[23] the *IGF1* 19-CA wild-type repeat status has been associated with higher IGF-I serum levels[13] than a nonwild-type repeat status, and higher IGF-I serum levels have been associated with an increased CRC risk.[24] For comparison, in the study by Simons *et al.*,[21] a decreased CRC risk in women was found for *IGF1* 19-CA variant *vs.* wild-type repeat carriers regardless of BMI. Although this is not entirely consistent with the present findings, one could carefully speculate that the CRC protective effects of *IGF1* 19-CA variant repeat alleles might be negated by a metabolically unhealthy state of prolonged exposure to IGFs and other growth factors, which may be captured by the presence of T2DM but not BMI, at least not in the NLCS. The NLCS includes relatively few individuals that reported to be obese (BMI > 30 kg/m^2) at baseline (9.6%), which is consistent with the relatively low prevalence of diabetes. T2DM, in this regard, as opposed to BMI, may have better or very specifically captured metabolically unhealthy individuals in our cohort.

Strengths of our study include the prospective design, limiting the chance of selection and information bias, the fact that we had information on many potential confounders, and a large number of CRC cases. In addition, the use of a GRS reduced the number of tests that had to be performed, minimizing the chance of false positive results due to multiple testing. However, the prevalence of T2DM in this cohort was low, limiting power. Another limitation of our study is that we had no repeated exposure measurements during the long follow-up period. As such, HRs may have been underestimated, considering that it is likely that some of the nondiabetic individuals were diagnosed with T2DM after our baseline measurement, as the population grew older. These individuals will have been misclassified as nondiabetic individuals in the present study, resulting in a smaller exposure contrast between the groups. Furthermore, classification of T2DM was performed on the basis of self-reports, which could have resulted in underdiagnosed T2DM in our cohort. However, previous studies within large cohorts have shown that the self-report of treated diabetes is an accurate proxy to use in epidemiologic studies.[25]

To conclude, our study of joint effects indicated that the presence of unfavorable alleles in the IGF pathway may further increase the risk of CRC associated with T2DM. Genetic variation in the IGF pathway could prove useful in future studies assessing personalized CRC risk.

Acknowledgement

This independent work was financially supported by the Dutch Cancer Society (grant number 2009-4281 to MPW) and the Health Foundation Limburg.

References

1. *GLOBOCAN Cancer Fact Sheets: Colorectal Cancer.* Available from: http://globocan.iarc.fr/factsheet.asp [date last accessed March 27, 2013].
2. Danaei G, Finucane MM, Lu Y, et al. National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. *Lancet* 2011;378:31–40.
3. Tsilidis KK, Kasimis JC, Lopez DS, et al. Type 2 diabetes and cancer: umbrella review of meta-analyses of observational studies. *BMJ* 2015;350:g7607.
4. Giovannucci E. Insulin, insulin-like growth factors and colon cancer: a review of the evidence. *J Nutr* 2001;131:3109S–20S.
5. Simons CCJM, Schouten LJ, Godschalk RWL, et al. Genetic variants in the insulin-like growth factor pathway and colorectal cancer risk in The Netherlands cohort study. *Sci Rep* 2015;5:14126.
6. de Kort S, Simons CCJM, van den Brandt PA, et al. Diabetes mellitus type 2 and subsite-specific colorectal cancer risk in men and women: results from The Netherlands cohort study on diet and cancer. *Eur J Gastroenterol Hepatol* 2016;28:896–903.
7. Hu FB, Manson JE, Liu S, et al. Prospective study of adult onset diabetes mellitus (type 2) and risk of colorectal cancer in women. *J Natl Cancer Inst* 1999;91:542–7.
8. van den Brandt PA, Goldbohm RA, van 't Veer P, et al. A large-scale prospective cohort study on diet and cancer in The Netherlands. *J Clin Epidemiol* 1989;43:285–95.
9. van Breda SG, Hogervorst JG, Schouten LJ, et al. Toenails: an easily accessible and long-term stable

source of DNA for genetic analyses in large-scale epidemiological studies. *Clin Chem* 2007;53: 1168–70.
10. Casparie M, Tiebosch ATMG, Burger G, et al. Pathology databanking and biobanking in The Netherlands, a central role for PALGA, the nationwide histopathology and cytopathology data network and archive. *Cell Oncol* 2007;29:19–24.
11. Schouten LJ, Höppener P, van den Brandt PA, et al. Completeness of cancer registration in Limburg, The Netherlands. *Int J Epidemiol* 1993;22:369–76.
12. Limburg PJ, Anderson KE, Johnson TW, et al. Diabetes mellitus and subsite-specific colorectal cancer risks in the Iowa Women's health study. *Cancer Epidemiol Biomark Prev* 2005;14:133–7.
13. Rosen CJ, Kurland ES, Vereault D, et al. Association between serum insulin growth factor-I (IGF-I) and a simple sequence repeat in IGF-I gene: implications for genetic studies of bone mineral density. *J Clin Endocrinol Metab* 1998;83:2286–90.
14. Barlow WE. Robust variance estimation for the case-cohort design. *Biometrics* 1994;50:1064–72.
15. Hughes LAE, Schouten LJ, Goldbohm RA, et al. Self-reported clothing size as a proxy measure for body size. *Epidemiology* 2009;20:673–6.
16. Rothman KJ, Greenland S. *Modern epidemiology*. Philadelphia, PA: Lippincott-Raven Publishers, 1998.
17. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Stat Med* 2000;19:1141–64.
18. Campbell PT, Deka A, Jacobs EJ, et al. Prospective study reveals associations between colorectal cancer and type 2 diabetes mellitus or insulin use in men. *Gastroenterology* 2010;139:1138–46.
19. de Kort S, Masclee AAM, Sanduleanu S, et al. Higher risk of colorectal cancer in patients with newly diagnosed diabetes mellitus before the age of colorectal cancer screening initiation. *Sci Rep* 2017;7:46527.
20. Olivo-Marston SE, Hursting SD, Lavigne J, et al. Genetic reduction of circulating insulin-like growth factor-1 inhibits azoxymethane-induced colon tumorigenesis in mice. *Mol Carcinog* 2009;48:1071–6.
21. Simons CCJM, Schouten LJ, Godschalk R, et al. Body size, physical activity, genetic variants in the insulin-like growth factor pathway and colorectal cancer risk. *Carcinogenesis* 2015;36:971–81.
22. Yuhara H, Steinmaus C, Cohen SE, et al. Is diabetes mellitus an independent risk factor for colon cancer and rectal cancer? *Am J Gastroenterol* 2011;106:1911–21. quiz 1922.
23. Keku TO, Vidal A, Oliver S, et al. Genetic variants in IGF-I, IGF-II, IGFBP-3, and adiponectin genes and colon cancer risk in African Americans and whites. *Cancer Causes Control* 2012;23: 1127–38.
24. Kaaks R, Toniolo P, Akhmedkhanov A, et al. Serum C-peptide, insulin-like growth factor (IGF)-I, IGF-binding proteins, and colorectal cancer risk in women. *J Natl Cancer Inst* 2000;92: 1592–600.
25. Jackson JM, Defor TA, Crain AL, et al. Self-reported diabetes is a valid outcome in pragmatic clinical trials and observational studies. *J Clin Epidemiol* 2013;66:349–50.